Information Sciences 264 (2014) 19–31

Contents lists available at ScienceDirect

Information Sciences

journal homepage: www.elsevier.com/locate/ins

Improving arithmetic skills through gameplay: Assessment of the effectiveness of an educational game in terms of cognitive and affective learning outcomes

Elena Núñez Castellar [a,*], Jan Van Looy [a], Arnaud Szmalec [b], Lieven de Marez [a]

[a] Department of Communication Sciences, iMinds-MICT, Ghent University, Belgium
[b] Psychological Sciences Research Institute, Université Catholique de Louvain, Belgium

ARTICLE INFO

Article history:
Available online 19 September 2013

Keywords:
Arithmetic training
Mental calculation
Educational game
Effectiveness
Cognitive learning outcome
Affective learning outcome

ABSTRACT

The present study assesses the effectiveness of a commercial educational math game for improving the arithmetic skills of children. Eighty-eight second graders were randomly assigned to one of three groups: a 'gaming group' which was instructed to play through the entire commercial game 'Monkey Tales', a group which was instructed to complete math exercises on paper and a control group that that did not receive any arithmetic exercises. We used a multidimensional approach to estimate the impact of game playing on objective measures of arithmetic performance such as speed and accuracy on a math test, as well as subjective measures such as math anxiety, enjoyment and perceived competence. Overall, the present study shows that the use of games for arithmetic can be beneficial both in terms of affective and cognitive learning outcomes.

1. Introduction

Playing digital games has become a popular pastime amongst children. According to the Federation of American Scientists, children aged 8–18 spend on average 50 min per day playing digital games [20]. In Europe, the statistics of digital games use show a similar trend. According to the EU Kids Online network, 9–16 year old internet users spend 88 min per day online and, when children start using the internet, playing games is reported as the second most common activity after finishing schoolwork [14].

As digital games increase in popularity, game developers realized the potential of capitalizing on their entertainment value by offering instructional content during game play [22]. This has given rise to the genre of "serious games" that employ the medium's rich, role-playing, story-based environments to teach, train, and change knowledge, attitudes, and behavior [7]. While serious games have been used in a broad range of domains, e.g. military, government, education, corporate, healthcare, the present study focuses on the use of math games in education.

1.1. Math games in education

The idea of using games to help children learn math is not new. According to Van Eck and Dempsey [26], in 1985 a large study was conducted by the National Council of Teachers of Mathematics (NCTM, United States), who tested eleven math

* Corresponding author. Address: Department of Communication Sciences, Ghent University, Korte meer 7-9-11, 9000 Ghent, Belgium. Tel.: +32 9 264 91 84.
E-mail address: elenapatricia.nunezcastellar@ugent.be (E. Núñez Castellar).

0020-0255/$ - see front matter © 2013 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.ins.2013.09.030

games for different grades with 1637 participants. The authors concluded that instructional games could be effective for higher grades when designed according to the curriculum and implemented with instructional activities.

To the best of our knowledge, the study conducted by the National Council of Teachers of Mathematics is the only study in which games for all grade levels were systematically tested. Most of the other studies reported in the literature focus on games that have been developed for a particular age group. Recently, it has been shown that very young children can also benefit from the use of math games. In a recent study, an internet game was used in conjunction with other classroom activities to teach probability and statistics to 4- and 5-year-old children. The children who played the game showed improvement in their understanding of probability [18]. Likewise, a recent study conducted with second graders (7–8 years old) showed that the use of a handheld game (Skills Arena [13]) was beneficial to students learning mathematics, in particular for low-ability students, regardless of gender and ethnic background [21]. The results demonstrated that pupils who played the game outperformed pupils who did not on a mathematics test. Likewise, the scores that the students obtained in the game correlated significantly with achievement scores on the mathematics test. Additionally, in a recent experimental study assigning 10 high school algebra and geometry classrooms to either a treatment or a control group has shown that students who played a math game (Dimension M [4]) improved significantly in mathematics performance compared to their non-gaming peers, but there were no significant improvements found in their motivation. Prior mathematics knowledge, computer skills, and English skills did not contribute to the students' motivation or mathematics performance [10]. Similar results suggesting that playing games can lead to gains in accuracy and calculation speed and improvement in attitude towards math in 10 to 11 years-olds, have also been reported [16]. All the studies cited above suggest that children benefit from playing games as compared to those who do not.

One crucial issue that has rarely been investigated is how using a game differs from using traditional methods. In other words, to what extent does a gaming environment add value to math education, over and above the learning effects of traditional methods? Without a control condition involving classical training methods, it cannot be ruled out that the beneficial effects of playing an educational math game are not bound to the gaming environment but merely to the exposure to mathematical content. To our knowledge, only two studies have touched upon this issue. Koran and McLaughlin [11] compared the effectiveness of drill and a mathematics game in teaching basic multiplication to 5th graders randomly assigned to 2 groups. After a baseline phase of 6 instructional days, each group received either a math game activity or drill for 10 instructional days. The results showed that both activities were equally effective in teaching multiplications and improving math grades. However, between the two groups, the children who played the game reported more enjoyment than the children who used traditional methods. Similarly, an empirical study in which the outcomes of playing a math game vs. paper-and-pencil math exercises were tested, showed that 4th and 5th grade elementary school pupils developed more positive attitudes toward mathematics after five weeks of playing a computer math game (Astra Eagle [25]), but that there was no significant effect of computer gaming on students' cognitive test performance compared with paper and pencil drills [9]. While these studies provided some indications as to the effectiveness of using math games for older children, several questions remain regarding their effectiveness for younger children and the actual learning gains compared to paper exercises and no assignment. The objective in the current study is to assess the cognitive and affective learning outcomes of playing a commercial math game compared to paper math exercises and no assignment over the course of three weeks.

1.2. Assessment of the effectiveness of math games in education

As pointed out by de Freitas [3], the field of serious games needs more rigorous baseline studies that map the different uses of serious games and quantify their effectiveness. While there is currently a lack of consensus about how the effectiveness of serious games should be assessed, a general practice in the field is to consider not only the educational but also the motivational impact of games, as both cognitive and affective factors are known to interact in many aspects of daily life. According to McFarlane et al. [15] there are three areas in which the use of digital games can be effective in education: to train general cognitive abilities and skills, to evoke positive affective reactions and thus stimulate motivation, and for knowledge- and content-related learning. Consistent with this idea, the present study conceptualizes learning effectiveness of an educational game as a multidimensional concept. As suggested by Kraiger et al. [12], we assume that learning may be evident from changes in cognitive, affective or skill capacities. An important assumption of this approach is that these learning outcomes are not discrete but are usually interacting; changes in cognitive outcomes could for instance co-occur with changes in affective outcomes. Accordingly, evidence of the success of a specific training program may be derived from mean differences between pre- and posttest measures linked with anticipated learning outcomes. Changes in cognitive capacities can be assessed objectively using measures like amount of knowledge and accuracy and speed of recall (accessibility of knowledge). Affective outcomes such as attitudinal and motivational changes can be measured subjectively through self-report (for more details see [12]).

Accordingly, in the present study we combine objective with subjective measures to assess the cognitive and affective outcomes of an educational math game. In order to reach this goal, second graders were randomly assigned to one of three groups: a 'gaming group', a group which was instructed to complete a paper exercises and a control group that that did not receive any mathematical assignment. The experimental design including three groups instead of two (control vs. gaming), constitutes an important methodological aspect of the present study. By comparing paper drills to game playing relative to control condition, we are able to experimentally isolate the effects of the "gaming experience" and thus estimate whether the gaming experience improves arithmetic performance over and above the training effects expected from the classical methods. To the best of our knowledge, there are no published studies in which the same methodology – using three groups for comparison – has been used to investigate the effectiveness of math games.

Based on the literature [9,11] we hypothesize that playing the educational game will lead to significant changes in affective learning as compared to the drill exercises but that they will not differ in terms of cognitive outcomes. More specifically we expect that the educational game, in terms of accuracy and speed, will be as effective as paper drills in improving the arithmetic skills of children. However, we hypothesize that the educational game will lead to greater anxiety reduction, a first affective learning outcome, than the paper math drill exercises or the control condition. Furthermore, regarding enjoyment and perceived competence, two more affective learning outcomes, we expect that the self-reported scores for these two variables will be significantly higher in the gaming group than those in the paper exercises group.

2. Method

2.1. Design

Children were randomly assigned to three groups: a gaming group, a paper exercises group, and a control group. The gaming group was instructed to play through the entire educational game Monkey Tales in three weeks' time. The paper exercises group was instructed to complete a set of math drill exercises in the same period, equivalent in quantity and basic level of difficulty to the exercises in Monkey Tales. The control group did not receive any assignment. Children were tested at two points over a three-week period: before (Pretesting) and after (Posttesting).

2.2. Participants

Letters were sent to several schools in the area of Ghent, Belgium, to recruit participants for the study. The parents interested in having their children participate registered via the Computer-Aided Registration Tool for Experiments (CORTEX) [5].

In the first evaluation (pretest measurement), 88 second graders (58 boys and 30 girls) were tested. Parents gave written informed consent for their child's participation. At the second evaluation (posttest), only 84 children could be assessed (one child could not participate because of illness and three parents did not react to the repeated calls for posttest). From this sample, three participants were excluded because they were clinically diagnosed with disorders listed in the Diagnostic and Statistical manual of Mental disorders (DSM-IV) [1], namely learning disability, ADHD, and dyslexia. In addition, participants who could not complete the task assignment were excluded (two children did not complete the game and one did not finish the paper math drills). Participants, who performed the computer math test at chance level or below, either in the pre- or the posttest, were excluded from the analyses, ensuring that all the participants included in the analyses were engaged in the task.

The data reported here includes 74 children. The participants' socio-demographic data is reported in Table 1. The groups did not differ significantly in terms of age, gender or game and study habits (see Table 1).

2.3. Stimulus material

2.3.1. Educational game

We used the 3D video game Monkey Tales [24], which exists in different versions for second to sixth grade and is used to support the learning of math. The game was developed with the active participation of teachers, schools, universities and educational method-developers based in the broadly used educational method of Die Keure, and has been adapted to the curricula of countries like the USA, UK and Belgium among others. In this study, we used the Belgian, Dutch-spoken version.

The main goal of this educational game is to improve the mental arithmetic skills of children by motivating them to engage in drill exercises with increasing time pressure. They can go through all the game levels only by being faster than a monkey (artificial intelligence). Importantly, the game uses an algorithm that tries to establish where a child is on the learning curve, and then stimulates the child to make progress by progressively augmenting the difficulty of the exercises. For the present study, we selected the Museum of Anything, which features math games for children in the 3rd grade (ages 8+). We chose this version given that the goal of the game is to train mental arithmetic and this is the version that is based on the second grade curriculum. Also at the time of testing all children were in the last months of the second grade (they were almost 3rd graders).

The educational game is divided into chapters and levels in which the player has to solve 3D puzzles and is challenged by a monkey to take part in a mini-game. The 3D puzzles require moving something that blocks the way or neutralizes a laser. The mini-game is a math exercise in classic game format (e.g. 2D shoot'em up). The player has to win the mini-game in order to go to the next level (see Fig. 1). The game contains 42 basic levels and 1 final level. Each level contains a mini-game, which includes an average of seven math exercises (e.g., $4 \times 4 = ?$). Only the final level contains mini-games. Appendix A, contains an overview of the mathematic contents of the mini-games. In order to complete all the levels of the game, children need to finish 322 math exercises whereby the exact number depends on how many times they need to replay a mini-game because of mistakes.

Table 1
Socio-demographic data and study and game habits by group.

	Educational game (N = 25)	Paper exercises (N = 23)	Control (N = 26)		
	n	n	n	Chi^2	p
Male gender	18	15	18	0.26	.88
	Mean	Mean	Mean	F	p
Age	7.52	7.26	7.35	1.33	.27
	Median	Median	Median	Chi^2	p[b]
Level education parents					
Education level father[a]	4	4	4	1.30	.52
Education level mother[a]	4	4	4	4.51	.10
	Mean	Mean	Mean	F	p
Study and game habits					
Homework hours per week	2 h 02 min	2 h 10 min	1 h 30 min	0.87	.42
Math homework hours per week	0 h 53 min	1 h 08 min	1 h 05 min	0.49	.61
Gaming hours during the week	3 h 41 min	3 h 42 min	3 h 11 min	0.48	.62
Gaming hours during the weekend	2 h 48 min	2 h 23 min	2 h 21 min	0.53	.59

[a] Four levels: Primary = 1, Junior High School/Middle School = 2, High School = 3, College/University = 4.
[b] Independent Sample Kruskall–Wallis Test.

2.3.2. Paper exercises

As one of the goals of the present study was to compare Monkey Tales with paper exercises, the latter needed to be as similar as possible to the former. Therefore, the educational publisher of the game, Die Keure, provided us with exercises based on their educational method that are equivalent to the basic level of difficulty of the exercises included in Monkey Tales. Of approximately 1000 exercises we received, a sample of 340 exercises that were representative of the Belgium math curriculum for second graders were selected. The exercises were organized in ascendant order of difficulty (as it is done in the educational game), and were given to the parents of the children in a folder that they gave back to us at the posttest. By checking this material, we verified that the children completed all the exercises.

2.4. Measures

2.4.1. Math performance: accuracy and speed

Publisher Die Keure (see tests A and B in Appendix B) provided two equivalent versions of exams (Test A and Test B) for assessing the math skills of children of the second grade. We used questions from these two tests to program a computerized version that allowed us to measure automatically the accuracy rates and the time-to-completion of the test in seconds. We programmed this computerized version using Tscope. Tscope is a C/C++ experiment-programming library for cognitive scientists. It provides functions for graphics, sound, timing, randomization and response registration [23]. All children performed the computer math test in the pre- and post-session. In each group, half of the children performed Test A as pretest measurement and Test B as posttest measurement. The other half performed the tests in the opposite order. As found in the appendix both tests are equivalent.

2.4.2. Math anxiety

In order to assess changes in math anxiety, we applied the Scale for Early Mathematics Anxiety (SEMA) [27] in the pre- and posttest sessions. The SEMA is a 20-items questionnaire that has shown to be a reliable and construct-valid measure of math anxiety in 7 to 9-year-old second and third graders.

Since the scale was only available in English, we translated it following a standard procedure to ensure the validity of the scale after the translation. Briefly, the following steps were followed for the translation. First, a bilingual, native speaker of Dutch translated the scale. Afterwards three reviewers who studied Germanic languages/ and or translation English–Dutch scored the translation of each item using a scale from 1 to 10. If the translation in Dutch was in their opinion perfect they scored the item with 10 (e.g. Is this right?: 9 + 7 = 18/Klopt dit?: 9 + 7 = 18). Similarly, if the reviewer would change one or two words but overall the meaning of the translation was correct then 9 was used as score. However, if the reviewer changed more than that, then the translation was scored below 9 and they were asked to add a suggestion (indicate what they would change to improve it). Finally, based on these suggestions we modified the problematic items (items scored with 8 or less), to ensure that the average score for all items was 9 or above. The result of the Dutch translation can be found in Appendix C. The administration of the SEMA was done as described by Wu et al. [27]. The individual SEMA scores were computed by adding up the 20 items' ratings. The range between the minimum and maximum SEMA scores is 20–100.

2.4.3. Enjoyment

In the second session, children filled out a questionnaire aimed to assess their enjoyment playing the educational game or solving the paper exercises. The items were taken from the extended short feedback questionnaire [17]. The first two questions

Fig. 1. Screen shoot Monkey Tales.

aimed to investigate enjoyment and perceived competence. Participants were asked to indicate in the 'funometer' [19] how much they enjoyed and whether they were good at playing or solving the paper exercises. The third question aimed to evaluate how children would describe their experience. Children were asked to select one or more attributes that describe their experience of playing the educational game or solving the paper exercises. The attributes included were "great", "tiring", "boring", "confusing", "exciting", "fun", "difficult", "intuitive", "simple" and "childish" (accordingly in Dutch: "fantastisch", "vermoeiend", "saai", "verwarrend", "spannend", "plezierig", "moeilijk", "intuïtief", "eenvoudig" and "overdreven kinderachtig"). Finally, children were asked whether they would like to play the educational game or do the math exercises again.

2.5. Procedure

The participants were tested at the beginning of May 2012 for the pretest session and at end of May for the posttest session. All pretest and posttest sessions were carried out at the Department of Communication Sciences, Ghent University. We used three rooms for testing: one for parents, one where children performed the computer math test and one for the math anxiety and enjoyment tests. Accordingly, three researchers were assigned to each room to supervise the tests.

As previously described, children were randomly assigned to three groups. One group of children was instructed to finish the educational game Monkey Tales in three weeks' time (Gaming group). Parents were instructed to help with the software installation and support the children while playing the game tutorial. However, the parents were explicitly asked not to help children with the math exercises. Moreover, the parents were asked to monitor on a weekly basis the children progress in the game, and to motivate them to play if needed. Importantly the parents were briefed about how to check the progress and detect when children had completed all the levels of the game. Finally, one week before the posttest an e-mail was sent as a reminder that, by the end of the week, the children should have completed the game. Finally, in the posttest session parents as well as children were asked apart whether the game was completed.

During the same three weeks' time, a second group of children was instructed to complete a set of math drill exercises on paper of equivalent quantity and basic level of difficulty as the exercises in Monkey Tales (Paper exercises group). Similar to the gaming group, the parents of the paper exercise group were instructed to check on a weekly basis and motivate the children to do the math drill exercises, but not to help them. One week before the posttest an e-mail was sent as a reminder that by the end of the week children should have completed the math drill exercises.

The control group did not receive any assignment but served as a comparison. The parents of this group were asked not to change anything from their normal routine and specifically they were asked not to let their children play any educational math games.

All parents received the instruction to let the children continue to do their math homework as usual. The group that completed the paper exercises and the control group received the educational game at the end of the posttest as a reward. All the parents received 15 euros for their participation.

2.6. Data analyses

First, we calculated descriptive statistics of the socio-demographic variables of each group. This to control that the random allocation of individuals to each group had not involuntarily resulted in systematically biased groups. In addition, the scores obtained from all learning outcome measures at the time of the pretest were compared between groups, using analysis of variance (ANOVA), to control possible baseline differences between the groups before assignment to the experimental

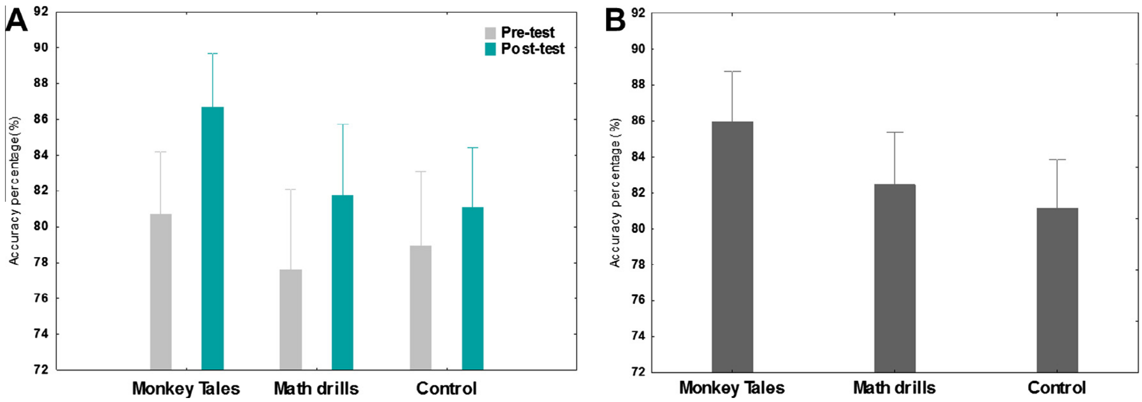

Fig. 2. (A) Mean accuracy rates in the pre and posttest measurements by group. (B) Mean accuracy rates in the posttest measurement, adjusted by pretest performance.

conditions. Having controlled for this, the next step was to consider the posttest scores of the three groups using the pretest performance scores as a comparative baseline.

In order to estimate possible differences in performance between the pre- and the posttest, the data obtained from the computer math test and the math anxiety ratings were analyzed using an analysis of covariance (ANCOVA) on the posttest scores, with the pretest scores entered as a covariate. Importantly analyses of covariance were only carried out when the assumptions of the ANCOVA model [6] were met. The variable Group (Gaming vs. Paper exercises vs. Control) was used as a between-subjects factor and the pretest scores were entered as covariates. Subsequently, post hoc tests were performed based on Tukey's HSD (Honestly Significant Difference), to correct for multiple comparisons. For the analyses of the self-report measures – aimed to assess enjoyment and perceived competence – we first calculated descriptive statistics. The results revealed the distribution of the scores were significantly skewed and significantly non-normal in terms of kurtosis. Significance was determined by comparing the numerical value for skewness and kurtosis with twice the respective standard error for both measures (Enjoyment: skewness = −2.01, kurtosis = 3.65, Perceived competence: skewness = −0.68, kurtosis = 1.38). Considering this we decided to recode the enjoyment and perceived competence scores into dummy variables (high vs. low scores) and by means of a chi-square test of independence, to investigate whether there were significant differences between groups. For all parametric and non-parametric tests, a significance level of 0.05 was used.

3. Results

3.1. Math performance

3.1.1. Accuracy rates

A one-way between-groups analysis of covariance was conducted on the posttest accuracy scores using pretest scores as a covariate (see Fig. 2B). The results indicated that after the influence of the pretest was controlled, the posttest scores of the three groups were significantly different ($F(2.70) = 3.20$, $p < 0.05$). Post hoc comparison using the Tukey test showed that the posttest accuracy of the group that played the educational game was significantly better than the accuracy of the paper exercises group ($p < 0.05$) and of the control group ($p < 0.05$). This indicated that the educational game had a stronger learning effect on the students' accuracy in the computer math test when compared with the group of children who did the paper math drills and the group that received no assignment. The post hoc test revealed that in terms of accuracy the adjusted average scores of the paper exercises group was not better than the scores of the control group ($p = .93$), indicating that only the group that played the educational game showed gain in the posttest scores when compared with the other groups.

3.1.2. Time-to-completion

The time-to-completion of the math test in the posttest was subject to a one-way between-groups analysis of covariance using the pretest times as a covariate (see Fig. 3). The results indicated that after the influence of the pretest was controlled, the adjusted means of the posttest differed significantly between groups ($F(2.70) = 3.78$, $p < 0.05$). Post hoc Tukey comparisons indicated that the posttest time-to-completion of the group that played the educational game was significantly better than that of the control group ($p < 0.05$) but not than the time-to-completion required by the paper exercises group ($p = 0.85$). This indicated that the group of children that played the educational game only outperformed the control group and not the paper group in terms of speed of completion. The post hoc test also revealed that when comparing the time-to-completion of the paper math drills group and the control group, the group that did the paper math drills showed a trend to perform the computer math test faster than the control group ($p = .06$).

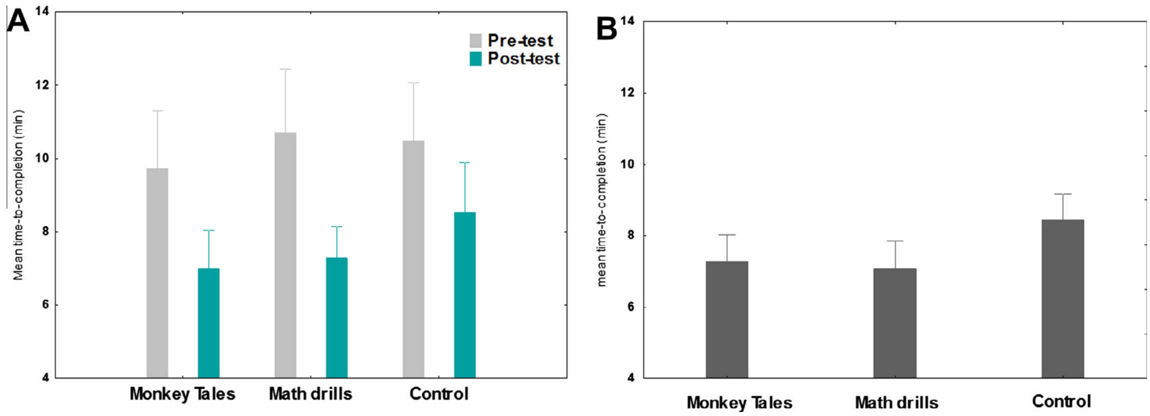

Fig. 3. (A) Mean time-to-completion in the pre and posttest measurements by group. (B) Mean time-to-completion in the posttest measurement, adjusted by pretest performance.

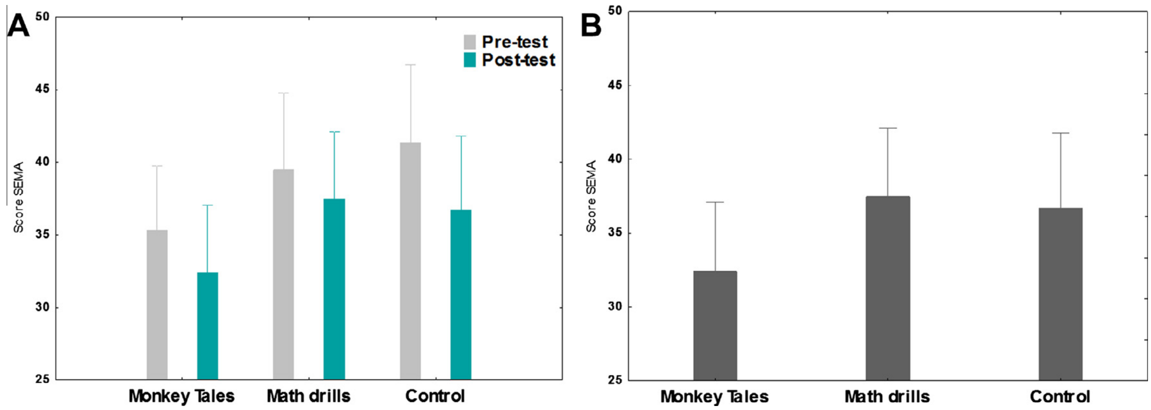

Fig. 4. (A) Mean math anxiety scores in the pre and posttest measurements by group. (B) Mean math anxiety scores in the posttest measurement, adjusted by pretest performance.

3.2. Math anxiety ratings

Math anxiety ratings were also subject to a one-way between-groups analysis of covariance using the pretest times as a covariate (see Fig. 4). The results indicated that after the influence of the pretest was controlled, the posttest scores of the three groups were not significantly different ($F(2.70) = 3.20$, $p = .50$). This indicates that, in terms of math anxiety ratings, we failed to find evidence suggesting that after playing the educational game participants showed a math anxiety reduction.

3.3. Enjoyment

3.3.1. Enjoyment and perceived competence

A chi-square test of independence was performed to examine the relation between enjoyment and the experience of gaming or doing the paper exercises. The results indicate a statistically significant association, χ^2 (1, $N = 48$) = 4.74, $p < .05$. Children were more likely to report high enjoyment when playing the game than when doing paper exercises. Additionally a chi-square test of independence was conducted to examine the relation between perceived competence and the experience of gaming or doing the paper exercises. The results did not indicate a statistically significant association, χ^2 (1, $N = 48$) = 0.04, $p = .95$, between the perceived competence of children for playing the game or doing the paper exercises.

3.3.2. Describing the game/math drills experience

Children were asked to select one or more attributes that describe their experience of playing the educational game or solving the paper exercises. The results per group for each of the attributes can be found in Fig. 5. The results of a chi-square test revealed that the frequency at which the attributes "boring", "fun", 'exciting', 'great' and 'simple' were selected differed significantly between the groups. A higher proportion of children selected the attribute "boring" to describe their experience solving the paper math exercises than to describe their experience playing the educational game, χ^2 (1, $N = 48$) = 6.07,

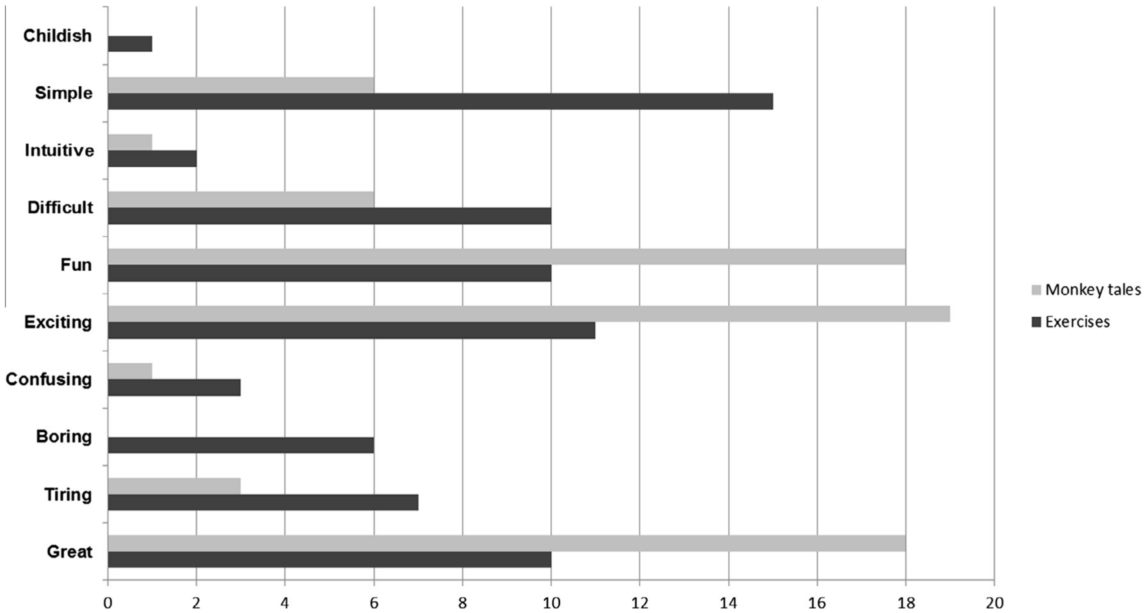

Fig. 5. Attributes used by children to describe their experience of playing the educational game or solving the paper exercises.

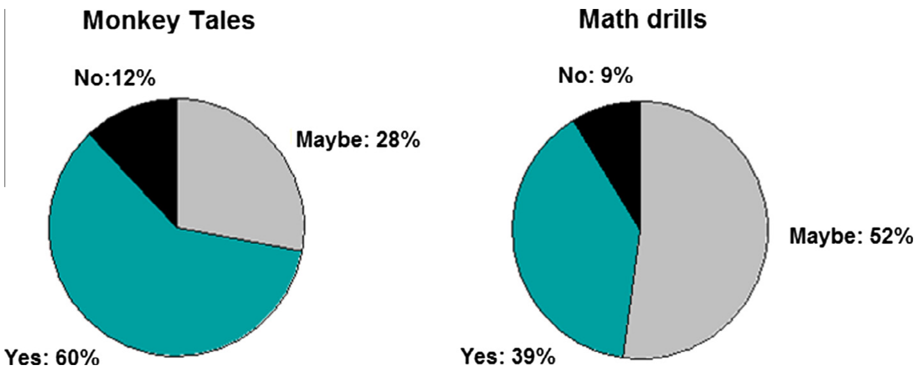

Fig. 6. Proportion of children who would like to play the educational game or do more math exercises again.

$p < .05$. Likewise, more children selected the attribute "fun" to describe their experience playing the educational game as compared to solving the paper math drills, χ^2 (1, $N = 48$) = 6.76, $p < .05$. The results also revealed that children more frequently reported the attributes 'exciting', χ^2 (1, $N = 48$) = 4.22, $p < .05$, and 'great', χ^2 (1, $N = 48$) = 3.30, $p < .05$, to describe the experience of playing the educational game than to describe their experience solving the paper math exercises. Likewise the attribute 'simple', χ^2 (1, $N = 48$) = 4.00, $p < .05$, was more frequently used to describe their experience solving the paper math drills. For all the other attributes, no significant differences were found between groups.

3.3.3. Willingness to play the game/do math drills again

The last question included in the questionnaire, assessed to what extent children would like to play the educational game or do the math exercises again. The results of a chi-square test revealed that the proportion of participants who stated they would like to repeat the experience tended to be higher in the group that played the educational game (60%) than in the group that did the paper math drills (39%), χ^2 (1, $N = 48$) = 3.05, $p < .08$ (see Fig. 6).

4. Discussion

With the present study, we tested the effectiveness of an educational math game designed to help children improve their arithmetic skills. For this purpose, second graders were randomly assigned to one of three groups: a 'gaming group' which was instructed to complete the game Monkey Tales, a 'paper exercise' group which was instructed to complete a set of paper math exercises, and a 'control' group that did not receive any mathematical intervention. We assessed the impact of game playing on children's arithmetic skills (accuracy and speed, math anxiety, enjoyment and perceived competence).

Based on the literature we hypothesized that playing the educational game would lead to significant changes in the affective learning outcomes as compared with doing paper exercises. In line with this prediction, the enjoyment results showed that children reported that playing the game was more enjoyable and when they were spontaneously asked to select one or more attributes to describe their experience, a significantly higher proportion of the children who played the educational game selected positive attributes like "fun", "exciting" and "great" to describe their experience. Likewise, a significantly higher proportion of the children who filled out the paper exercises described their experience as "boring" and "simple". In the same vein, a higher proportion of the children who played the educational game reported that they would like to play again in the future compared to the children who did the exercises. These results are in line with previous research showing that, when comparing math games with paper exercises, the former elicit more positive affective responses [9,11].

However, when children were explicitly asked whether they were good at playing the educational game or solving the paper exercises no differences were found between groups. Since previous research has shown that the perceived competence may be one factor that determines whether or not trainees would apply the skills they have learned in the future [2] or an indicator of learning during training [12], we included this as a relevant affective learning outcome. However, we failed to find evidence suggesting that playing the educational game leads to higher perceived competence compared to doing paper exercises.

A third affective outcome that was assessed was the change in math anxiety between pre- and posttest. The results did not support our prediction that the educational game would lead to greater anxiety reduction than the paper exercises or no assignment. While we have translated the scale from English to Dutch, we do not believe this to have affected the content validity of the scale, given that in the present study we replicate the results of the study of Wu and colleagues [27], which indicate that the SEMA scale is a reliable and construct-valid measure of math anxiety in 7 to 9-year-old children. Specifically, in the paper of Wu and colleagues [27] the average SEMA scores reported for second graders was 34.65 with a standard deviation of 12.74 while in our study the mean score was 35.49 with a standard deviation of 11.65, which makes the results of both studies equivalent. Instead, we think that our findings could be driven by the small sample size of the present study. While visual inspection of graph 4B seems to point to a stronger reduction in the anxiety scores for the Monkey Tales group, there is too much variance for the difference to be significant. Future research could investigate whether significant changes in anxiety scores can be found when testing larger samples or when using other, objective measurements like psychophysiological indexes of arousal to measure math anxiety.

Besides the affective learning outcomes, cognitive learning outcomes were also assessed in the present study. Two objective measurements – accuracy and time to completion – were considered to measure the impact of the educational game on the arithmetic skills of children. Specifically, we predicted that the educational game, in terms of accuracy and speed, would be as effective as paper math drill exercises in improving the arithmetic skills of children, and both groups would perform better than the control group. In line with this prediction the posttest results show that the group that played the educational game was significantly faster at completing the math computer test than the control group, but not faster than the paper exercises group. These results are in line with previous studies showing that, in terms of cognitive learning outcomes, math games yield similar results as paper drill exercises [9,11] and with the findings of Miller and Robertson [16], who observed that in terms of speed children who played a math game over a 9-week period were twice faster than controls in solving arithmetic problems after the training phase.

However, an unexpected finding occurred; the accuracy gain in the computer math test was significantly higher for the group of children who played the educational game. More specifically, the results showed that the accuracy of the group that played the educational game was significantly higher than that of the paper exercises and the control group, indicating that the educational game had a higher learning impact. There are three possible explanations for this phenomenon, none of which excludes the other. First, it is possible that, as children were more motivated to play the game, they engaged in additional self-training to improve their playing skills. A second possible explanation for the higher performance gain of the children who played the educational game is that, while the basic difficulty level was the same, they were offered specific exercises more in line with their math skills due to the adaptive algorithm in the game. This algorithm monitors the results of the player and changes the difficulty level (but not the type of exercise) of the math exercises accordingly, mainly by using higher values. This explanation is supported by the fact that, when choosing attributes to describe their experiences, the children doing the paper exercises described the activity significantly more often as "easy".

Finally an important element of the game, which is absent on the paper exercises, is the competitive element. In the game, the math exercises are competitively performed by the player against a monkey. Only by being faster than the monkey can children go to the next level. As previous research has shown, performing games in interpersonal competitive contexts can enhance arithmetic learning [8]. It is plausible that in the present study the competitive context has led to additional performance gains in terms of accuracy. Based on the present study it is not possible to determine which of the factors mentioned above caused or contributed most to the children reaching a higher accuracy percentage. Future research could explore the specific mechanism or mechanisms underlying this accuracy improvement.

Finally, one of the limitations of the present study is that, since the subject pool was self-selected (parents were free to contact us to participate in the study), selfselection might be a limitation of the present study. Future research should examine whether or not our findings can be generalized. Another limitation is that although we closely matched the number and difficulty level of the paper exercises, we did not get any detailed information about the exact number of exercises performed by children who played the educational game. We anticipated that children would fail about four mini-games so we added the same number of extra exercises on paper. However, it would be useful in future research to log the children's behavior while playing the game. Log data would allow researchers to measure the exact number of exercises performed during game play. In the present study, due to technical failure, we were unable to control for these factors. Future research could analyze actual in-game playing behavior and mental arithmetic performance and link them with learning outcomes.

5. Conclusion

To conclude, the results of the present study indicate that playing an educational game to help second grade pupils improve their arithmetic skills, can increase their accuracy compared to paper exercises or no exercise. Moreover, the math games can increase mental calculation speed in a similar way as an equivalent number of paper exercises. Furthermore, in terms of affective outcomes, playing the game was reported to be more enjoyable than filling out the paper exercises. We also assessed whether playing the educational game would lead to anxiety reduction, but failed to find evidence supporting this hypothesis.

In the present study we show that the extent to which children learn from an educational game may be evident from changes in cognitive and affective measurements in line with the proposal of Kraiger et al. [12]. Rather than concluding that a game is effective or not, this approach allows us to provide a more balanced picture to teachers and parents showing in which areas children can be expected to benefit from the use of educational games to train mental arithmetic.

Contributors

Elena Nunez Castellar, Jan Van Looy, Arnaud Szmalec and Lieven de Marez were involved in the design of the study. Elena Núñez Castellar and Jan Van Looy recruited the participants and managed the data collection. Jan Van Looy and Arnaud Szmalec supervised the statistical analyses and the manuscript preparation. Elena Núñez Castellar conducted the literature searches, ran the statistical analyses and wrote the manuscript. All authors contributed to and have approved the final manuscript.

Acknowledgements

This work was performed in the context of the IBBT Games at School project, which is cofunded by iMinds, a research institute founded by the Flemish Government. We thank Vicky Vermeulen and Nic Pappijn (Die Keure) the for their support with the materials and information about the educational game and Anissa All for helping us contacting the schools.

Appendix A. List of the material covered in the Monkey Tales game "The Museum of Anything"

Tens and units
Understanding times/multiplied by
Understanding tenths
Understanding hundreds
Even or odd up to 100
Divide into equal parts
Multiplication tables from 2 up to 10
Division tables from 2 up to 10
Understanding numbers up to 100
Addition with 1 multiple of 10 up to 100
Subtraction with 1 multiple of 10 up to 100
Addition and subtraction with 1 multiple of 10 up to 100
Addition with up to 3 multiples of 10 up to 100
Subtraction with up to 3 multiples of 10 up to 100
Addition and subtraction with up to 3 multiples of 10 up to 100
Addition and subtraction up to 100 without regrouping
Addition TU+U up to 100 with regrouping
Subtraction TU+U up to 100 with regrouping
Addition and subtraction TU+U up to 100 with regrouping
Addition TU+TU up to 100 with regrouping
Subtraction TU+TU up to 100 with regrouping
Addition and subtraction TU+TU up to 100 with regrouping

Appendix B. Equivalent versions of math exams – Test A and Test B

	Item	Problem	A	B	C	D
Test A						
1	Addition and subtraction up to 20 with regrouping	8 + 7 =	13	15	17	14
2	Greater and smaller up to 20	What number is 5 smaller than 12?	7	17	8	16
3	Half up to 20	What is 8 halved?	4	16	14	2
4	Double up to 20	What is 6 doubled?	3	14	18	12
5	Precedes up to 20	What number comes just before 17?	6	18	16	15
6	Follows up to 20	What number comes just after 13?	12	13	14	15
7	Patterns with leaps of 2 or 5	Complete: 6 8 10 .	12	14	10	16
8	Multiplications/divisions by 2	$9 \times 2 =$	18	8	16	14
9	Multiplications/divisions by 3	24:3 =	5	7	8	6
10	Multiplications/divisions by 4	$8 \times 4 =$	32	28	24	36
11	Multiplications/divisions by 5	45:5 =	9	8	7	6
12	Multiplications/divisions by 6	$6 \times 6 =$	36	32	54	42
13	Multiplications/divisions by 7	49:7 =	9	5	7	8
14	Multiplications/divisions by 8	$8 \times 8 =$	56	64	72	48
15	Multiplications/divisions by 9	72:9 =	7	8	6	9
16	Multiplications/divisions by 10	$4 \times 10 =$	40	50	30	60
17	Addition with 1 T up to 100	54 + 10 =	44	55	64	74
18	Subtraction with 1 T up to 100	86 − 10 =	76	85	66	96
19	Subtraction TE − TE = T	45 − 25 =	10	30	40	20
20	Addition up to 100 without regrouping	81 + 8 =	88	89	87	90
21	Subtraction up to 100 without regrouping	67 − 5 =	65	62	72	63
22	Addition up to 100 with regrouping TE + E	88 + 9 =	97	98	79	96
23	Addition up to 100 with regrouping TE + E	64 + 8 =	70	71	72	74
24	Subtraction up to 100 with regrouping TE − E	63 − 6 =	57	75	56	69
25	Subtraction up to 100 with regrouping TE − E	45 − 8 =	53	38	37	33
26	Tens and units	What number is 7 E?	76	17	70	71
27	Tens and units	What number is 5 T?	15	51	65	95
28	Knowledge of numbers up to 100	Complete: 30 40 50 .	40	30	60	70
29	Knowledge of numbers up to 100	Complete: 78 . 82 84	76	86	79	80
30	Addition with 3 T up to 100	56 + 30 =	66	86	96	76
31	Subtraction with 3 T up to 100	75 − 30 =	45	35	65	55
32	Structure of 100	100 = 20 + 20 + 20 + 20 + .	10	40	20	100
33	Even/uneven up to 100	Indicate the uneven number in the series.	16	61	72	44
34	Structure of tens	90 is 50 and .	30	20	40	10
35	Dividing into equal parts	What number fits? 16 = . + . + . + .	3	4	5	6
36	Addition up to 100 with regrouping TE + TE	57 + 24 =	80	71	81	77
37	Addition up to 100 with regrouping TE + TE	26 + 48 =	72	74	64	84
38	Addition up to 100 with regrouping TE + TE	55 + 29 =	84	74	85	86
39	Subtraction up to 100 with regrouping TE − TE	74 − 56 =	18	16	26	28
40	Subtraction up to 100 with regrouping TE − TE	51 − 38 =	24	23	14	13
Test B						
1	Addition and subtraction up to 20 with regrouping	5 + 8 =	13	15	12	14
2	Greater and smaller up to 20	What number is 7 smaller than 17?	7	11	10	9
3	Half up to 20	What is 10 halved?	5	15	20	2
4	Double up to 20	What is 4 doubled?	2	6	8	10
5	Precedes up to 20	What number comes just before 10?	11	12	8	9
6	Follows up to 20	What number comes just after 16?	17	15	14	18
7	Patters with leaps of 2 or 5	Complete: 0 5 10 .	20	15	25	10
8	Multiplications/divisions by 2	$6 \times 2 =$	18	12	10	10
9	Multiplications/divisions by 3	18:3 =	5	7	8	6
10	Multiplications/divisions by 4	$9 \times 4 =$	32	28	24	36

(*continued on next page*)

Appendix B. (*continued*)

	Item	Problem	A	B	C	D
11	Multiplications/divisions by 5	35:5 =	9	8	7	6
12	Multiplications/divisions by 6	9×6 =	36	32	54	42
13	Multiplications/divisions by 7	63:7 =	9	5	7	8
14	Multiplications/divisions by 8	7×8 =	56	64	72	48
15	Multiplications/divisions by 9	54:9 =	7	8	6	9
16	Multiplications/divisions by 10	0×10 =	0	50	100	60
17	Addition with 1 T up to 100	73 + 10 =	83	74	63	74
18	Subtraction with 1 T up to 100	63 − 10 =	53	73	62	43
19	Subtraction TE − TE = T	73 − 43 =	10	30	40	20
20	Addition up to 100 without regrouping	52 + 7 =	59	45	58	61
21	Subtraction up to 100 without regrouping	89 − 7 =	83	81	82	80
22	Addition up to 100 with regrouping TE + E	65 + 8 =	72	83	71	73
23	Addition up to 100 with regrouping TE + E	75 + 7 =	72	81	73	82
24	Subtraction up to 100 with regrouping TE − E	75 − 9 =	55	65	66	84
25	Subtraction up to 100 with regrouping TE − E	93 − 6 =	88	87	89	86
26	Tens and units	What number is 3 E?	73	37	30	31
27	Tens and units	What number is 7 T?	17	71	65	95
28	Knowledge of numbers up to 100	Complete: 40 50 60 .	40	30	60	70
29	Knowledge of numbers up to 100	Complete: 64 62 . 58	86	56	60	66
30	Addition with 3 T up to 100	47 + 30 =	77	17	67	57
31	Subtraction with 3 T up to 100	88 − 30 =	68	78	58	48
32	Structure of 100	100 = 40 + 40 + .	20	40	10	100
33	Even/uneven up to 100	Indicate the even number in the series.	17	31	72	99
34	Structure of tens	70 is 50 and .	30	20	40	10
35	Dividing into equal parts	What number fits? 21 = . +. + .	7	5	3	9
36	Addition up to 100 with regrouping TE + TE	46 + 37 =	82	84	93	83
37	Addition up to 100 with regrouping TE + TE	29 + 39 =	68	66	86	64
38	Addition up to 100 with regrouping TE + TE	37 + 45 =	82	83	72	92
39	Subtraction up to 100 with regrouping TE − TE	84 − 48 =	44	26	46	36
40	Subtraction up to 100 with regrouping TE − TE	75 − 56 =	19	20	29	21

Appendix C. Dutch translation SEMA

Instructies: “Nu ga ik je enkele wiskunde vragen laten zien. Ik wil dat je elke vraag leest en doet alsof je ze gaat beantwoorden. Dan wil ik dat je me vertelt hoe zenuwachtig het beantwoorden van deze vraag je maakt. Dus vergeet niet, je moet de vragen niet echt beantwoorden maar ik wil gewoon dat je doet alsof je deze gaat beantwoorden en kijkt welk gevoel je hier bij hebt. Het kan ervoor zorgen dat je je Helemaal niet zenuwachtig, Een klein beetje zenuwachtig, Redelijk zenuwachtig, Heel zenuwachtig of Heel erg zenuwachtig voelt. Begrijp je het?” (Oefening: Wie is de Eerste Minister van België?).

1. Jan kocht twee pizza’s die in zes stukken verdeeld waren. Hoeveel totale stuken moest Jan delen met zijn vrienden?
2. Klopt dit?: 9 + 7 = 18
3. Hoeveel geld heeft Ann als ze een halve euro en 20 cent heeft?
4. Hoe schrijf je het getal vierhonderdtweeëntachtig?
5. Teken een uren- en minutenwijzer op een klok, zodat zij 15u15 zou aangeven.
6. Teken een driehoek en een vierkant op het schoolbord.
7. Tel luid op per 5 van 10 tot 55.
8. Hoe laat zal het zijn binnen 20 minuten?
9. Klopt dit?: 15 − 7 = 8
10. Nele heeft meer geld dan Stijn. Stijn heeft meer geld dan Ineke. Wie heeft meer geld -Nele of Ineke?

Instructies: “Nu zal ik je een aantal zinnen voorlezen over situaties die te maken hebben met wiskunde. Probeer te doen alsof elke situatie echt gebeurt en denk erover hoe zenuwachtig ze je zou maken. Het kan ervoor zorgen dat je je Helemaal niet zenuwachtig, Een klein beetje zenuwachtig, Redelijk zenuwachtig, Heel zenuwachtig of Heel erg zenuwachtig voelt. Begrijp je het?” (Oefening: Je staat op het punt in een achtbaan te rijden).

1. Je bent in de rekenles en je leraar staat op het punt om iets nieuw aan te leren.
2. Je moet gaan zitten om aan je wiskundehuiswerk te beginnen.
3. Je bent al het geld in je spaarpot aan het optellen.
4. Iemand vroeg je een appeltaart te snijden in vier gelijke delen.
5. Je staat op het punt om een rekentest af te leggen.
6. Je bent in de rekenles en er is iets dat je niet begrijpt. Je vraagt je leraar om je te helpen.
7. Je leraar geeft je een hoop opteloefeningen om aan te werken.
8. Je leraar geeft je een hoop aftrekoefeningen om aan te werken. (uitleggen?)
9. Je bent in de klas bezig met een wiskundige oefening op het schoolbord.
10. Je luistert terwijl je leraar je uitlegt hoe je een rekenoefening moet oplossen.

References

[1] A. American Psychiatric, D.-I. American Psychiatric Association. Task Force on, Diagnostic and Statistical Manual of Mental Disorders: DSM-IV-TR, American Psychiatric Association, Washington, DC, 2000.
[2] A. Bandura, Self-efficacy determinants of anticipated fears and calamities, Journal of Personality and Social Psychology 45 (1983) 464–469.
[3] S. de Freitas, in: B. JISC (Ed.), Learning in immersive worlds, 2006.
[4] DimensionU, Dimension M, in: DimensionU, Inc. Long Island City, New York, 2010.
[5] M. Elson, G. Bente, in: CORTEX – Computer-Aided Registration Tool for Experiments, University of Cologne, Cologne, Germany, 2009.
[6] A. Field, Discovering Statistics Using SPSS for Windows, Sage Publications Ltd., London, 2000.
[7] R. Garris, R. Ahlers, J.E. Driskell, Games, motivation, and learning: a research and practice model, Simulation and Gaming 33 (2002) 441–467.
[8] F. Ke, B. Grabowski, Gameplaying for maths learning: cooperative or not?, British Journal of Educational Technology 38 (2007) 249–259
[9] F.F. Ke, A case study of computer gaming for math: engaged learning from gameplay?, Computers and Education 51 (2008) 1609–1620
[10] M. Kebritchi, A. Hirumi, H.Y. Bai, The effects of modern mathematics computer games on mathematics achievement and class motivation, Computers and Education 55 (2010) 427–443.
[11] L.J. Koran, T.F. McLaughlin, Games or drill: increasing the multiplication skills of students, Journal of Instructional Psychology 17 (1990) 222–230.
[12] K. Kraiger, J.K. Ford, E. Salas, Application of cognitive, skill-based, and affective theories of learning outcomes to new methods of training evaluation, Journal of Applied Psychology 78 (1993) 311–328.
[13] J. Lee, K. Luchini, B. Michael, E. Soloway, C. Norris, in: More than Just Fun And games: Assessing the Value of Educational Video Games in the Classroom, CHI 2004, 2004, pp. 1375–1378.
[14] S. Livingstone, L. Haddon, in: EU Kids Online: Final Report, LSE: EU Kids, 2009 (Online).
[15] A. McFarlane, A. Sparrowhawk, Y. Heald, in: Report on the Educational Use of Games: An Exploration by TEEM of the Contribution which Games can make to the Educational Process, TEEM, Cambridge, UK, 2002.
[16] D.J. Miller, D.P. Robertson, Educational benefits of using game consoles in a primary classroom: a randomised controlled trial, British Journal of Educational Technology 42 (2011) 850–864.
[17] C. Moser, V. Fuchsberger, M. Tscheligi, Rapid assessment of game experiences in public settings, in: Proceedings of the 4th International Conference on Fun and Games, ACM, Toulouse, France, 2012.
[18] J. Pange, Teaching probabilities and statistics to preschool children, Information Technology in Childhood Education Annual v2003 (2003) 163–172.
[19] J.C. Read, S. MacFarlane, Using the fun toolkit and other survey methods to gather opinions in child computer interaction, in: Proceedings of the 2006 Conference on Interaction Design and Children, ACM, Tampere, Finland, 2006.
[20] F.o.A. Scientists, Harnessing the Power of Video Games for Learning, in: Summit on Educational Games, 2006.
[21] N. Shin, C. Norris, E. Soloway, Effects of handheld games on students learning in mathematics, in: Proceedings of the 7th International Conference on Learning Sciences, International Society of the Learning Sciences, Bloomington, Indiana, 2006.
[22] T. Sitzmann, A meta-analytic examination of the instructional effectiveness of computer-based simulation games, Personnel Psychology 64 (2011) 489–528.
[23] M. Stevens, J. Lammertyn, F. Verbruggen, A. Vandierendonck, Tscope: AC library for programming cognitive experiments on the MS Windows platform, Behavior Research Methods 38 (2006) 280–286.
[24] L. Studios, Monkey tales, in: Die keure & Larian Studios Gent, Belgium, 2011.
[25] C.f.A. Technologies, Astra eagle, in: Center for Advanced Technologies, Altoona, PA, 2002.
[26] R. Van Eck, J. Dempsey, The effect of competition and contextualized advisement on the transfer of mathematics skills in a computer-based instructional simulation game, ETR&D-Educational Technology Research and Development 50 (2002) 23–41.
[27] S. Wu, H. Amin, M. Barth, V. Malcarne, V. Menon, Math anxiety in second and third graders and its relation to mathematics achievement, Frontiers in Psychology 3 (2012).