Improving arithmetic skills through an educational game

Elena Núñez Castellar[1], Jan Van Looy[1], Arnaud Szmalec[2], Lieven de Marez[1]

[1] Department of Communication Sciences, iMinds-MICT-Ghent University, Belgium

[2] Psychological Sciences Research Institute, Université Catholique de Louvain, Belgium

Corresponding author:

Elena Núñez Castellar

Department of Communication Sciences

Ghent University,

Korte meer 7-9-11, 9000, Ghent, Belgium.

Email: elenapatricia.nunezcastellar@ugent.be

Tel: +32 9 264 9184

Abstract

Abstract: Besides entertainment, games have shown to have the potential to impact a broader variety of cognitive abilities. Research has consistently shown that several aspects in cognition such as visual short-memory, multitasking and spatial cognition can be enhanced by game play. In a previous study, we found that playing Monkey Tales, a game aimed at training arithmetic skills, helped second grade pupils to increase their accuracy in mental calculation as compared to paper exercises (Nuñez Castellar, 2013). In this follow up study we explore whether traditional methods and game training differ in terms of the cognitive processes that both are able to impact. We incorporated standardized measures of working memory and visuo-motor skills. Additionally gains in arithmetical performance, and self-reported measures of enjoyment were investigated. We found some evidence suggesting that arithmetic performance enhancement induced by game play and paper exercises differ not only in terms of enjoyment but also of working memory capacity improvements.

Keywords: educational game, arithmetic training, working memory, visuo-motor skills, enjoyment, reaction times, accuracy.

Improving arithmetic skills through an educational game

Video games present themselves as one of the more interesting and promising means of improving cognitive abilities, particularly with children. One of their promises is that, compared to traditional training, they are more engaging and entertaining (Boot, Kramer, Simons, Fabiani, & Gratton, 2008). Moreover, besides entertainment, games have the potential to impact a broader variety of cognitive abilities. Recently, research has consistently shown that several aspects in cognition such as visual short-memory, multitasking and spatial cognition can be enhanced by game play (for a complete review, see Bavelier et al., 2012).

In a previous study (Núñez Castellar, Van Looy, Szmalec, & de Marez), we reported that playing Monkey Tales, a commercial game aimed at training arithmetic skills in children, helped second grade pupils to increase their accuracy in mental calculation as compared to paper exercises or no exercises. However, the extent to which the positive changes induced by gaming or by paper exercises in children differ in nature and extent is an issue that has not yet been explored. Specifically, based on previous research showing that video game playing can enhance working memory capacities and attention (Bavelier, Green, Pouget, & Schrater, 2012), in the present paper we explore whether, by incorporating standardized measures of working memory and visuo-motor skills before and after game training and training by means of math paper exercises, we can provide a more informed description of how arithmetic performance enhancement induced by these two methods might differ in terms of cognitive processes.

Specifically, there are reasons to believe that arithmetic performance enhancement induced by game play might be modulated by improvements in the domains of visuo-motor skills and working memory. Working memory is the ability to explicitly maintain a mental representation of a certain amount of information while being engaged simultaneously in other mental processes (Baddeley, 2000). Research has demonstrated that working memory capacity increases from preschool through the elementary school years. Preschool children can hold three to four items of information, such as numbers, in working memory, whereas a typical fourth grader can hold five to six items (Kail & Park, 1990). Although during the past decades it was traditionally assumed that working memory is highly heritable and unlikely to be influenced by environmental experience and opportunity (Campbell, Dollaghan, Needleman, & Janosky, 1997), recent findings have provided evidence suggesting that children's working memory can be enhanced by means of training (Klingberg et al., 2005; Turley-Ames & Whitfield, 2003). In fact, a recent review about the effects of cognitive training on children concluded that the training of core executive functions like working memory is most beneficial to 4-12 years-olds (Diamond & Lee, 2011).

Remarkably, a recent study of Holmes, Gathercole and Dunning (2009) has demonstrated that attention training can lead to a significant boost in the academic mathematics performance of children (Holmes, Gathercole, & Dunning, 2009). This study showed that IQ scores (both verbal IQ and performance IQ scores) did not show a comparable boost after working memory training, suggesting that, rather than leading to global performance enhancement, improvements in working memory seem to act locally, boosting arithmetical performance. Likewise, studies in cognitive psychology support this critical role of working memory. There is converging evidence showing that working memory capacity closely relates to skill in arithmetics and, in particular, to the speed of solving arithmetic problems (Geary & Widaman, 1992; Lemaire, Abdi, & Fayol, 1996; Rubinsten & Henik, 2009). Moreover, studies with clinical populations indicate the existence of a close relationship between working memory capacities and mathematical skills. For instance, studies investigating children with a mathematics learning disability (MD), have shown that they obtained diminished scores on a variety of working memory tasks when compared with their same age pairs (McLean & Hitch, 1999).

Visuo-motor integration is another cognitive ability that has been linked with mathematical achievements. For instance, research has provided evidence showing that children who have difficulties in math (aged 7 to 13 years) - but with normal reading skills - had a much higher frequency of poor performance in a test aimed to map visual-motor deficits (Developmental

Test of Visual-Motor Integration) (Siegel & Feldman, 1983). Likewise, research has demonstrated the existence of a close relationship between the Stanford total math standard score and the Developmental Test of Visual-Motor Integration when controlling for verbal ability and age (Sortor & Kulp, 2003). Also interestingly, a study that intended to predict reading and mathematics achievement in fourth-grade children from kindergarten scores in standardized tests, found that verbal skills uniquely predicted later reading achievement, whereas both verbal skills and visuo-motor skills uniquely predicted later mathematics achievement (Kurdek & Sinclair, 2001).

In addition, recent studies have reported evidence suggesting that working memory and visual-attention can be trained in normal adults by means of video gaming. For example, it has been found that video game players are faster and more accurate in the monitoring and updating of working memory than non-video game players (Colzato, van den Wildenberg, Zmigrod, & Hommel, 2012). Green and Bavelier (2003) conducted a series of experiments on the effects of video game playing on visual attention comparing action video game players and non-video game players, and found that video game playing experience enhances the capacity of the players' visual attention system (Green & Bavelier, 2003). Likewise several recent studies have demonstrated that action video game players have the ability to switch faster between tasks compared with non-video game players (Boot et al., 2008; Karle, Watter, & Shedden, 2010). Finally, a recent study has shown that performance gains are not restricted to the action game genre, but that playing Tetris, a casual puzzle game, can also improve working memory and visuo-spatial ability in young adults (Nouchi et al., 2013).

Taken together, the results mentioned above suggest the existence of a close link between, working memory, visuo-motor skills and arithmetic skills, and that, remarkably, these cognitive abilities can be trained by means of game play, especially with young children. This presents important opportunities for using games for mathematics training but also questions as to how these different performance gains are related and how they compare with traditional methods for practicing mathematics. Hence, in the present study, we explore whether traditional methods and game training differ in terms of cognitive processes that both are able to impact. In order to do this, we compared the results that second graders achieved in a test made for assessing their math skills and the scores of standardized measurements of working memory and visuo-motor skills before and after game training and traditional training by means of math paper exercises.

Standardized assessments of children's working memory, planning and visuo-motor skills were conducted before and after training by means of the Digit Span and Mazes subtests of the WISC-III NL which is a battery that provides a measure of IQ (Kort et al., 2002). The Digit span subtest measures the capacity to hold numbers in working memory and the ability to work with them. More specifically, the repetition of the digits (especially backwards) demands concurrent mental operations like divided attention, allocation of multiple mental resources operations, and active control of conscious attention (Pisoni, Kronenberger, Roman, Ann Geers, 2011). The subtest of Mazes measures not only visuo-motor abilities (Sattler,1988) but also forward planning and organization (Skuse,2003).

Finally, the present work also aimed to explore the relationship between gains in arithmetical performance (reaction times and accuracy), working memory and visuomotor skills and their relationship to the enjoyment of game training in comparison with traditional training. To the best of our knowledge, this is the first time that the predictive value of enjoyment is investigated in relation to gains in objective measures of arithmetic performance and cognitive abilities. Also on a methodological level it is the first time that a mathematics game was modified and its contents extracted to allow precise comparison between the gaming and paper exercises condition. Thus each single math exercise, type of question (e.g., multiple choice), quantity and order was perfectly matched in the game training and the traditional training.

Method

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) (Elson, 2009).

In the first evaluation (pretest measurement), 67 second graders (45 boys and 22 girls) were tested. Parents gave written informed consent for their child's participation. At the second evaluation (posttest), only 63 children could be assessed (three children didn't completed their assignment - either they didn't liked the game or had no interest in completing the paper drills- and one parent did not react to the repeated calls for posttest). From this sample, seven participants were excluded because they were clinically diagnosed with disorders listed in the Diagnostic and Statistical manual of Mental disorders (DSM-IV) (American Psychiatric & American Psychiatric Association. Task Force on, 2000), namely learning disability, ADHD, and dyslexia. In addition, four 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 52 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).

Insert Table 1 about here

Design

Children were randomly assigned to two groups. One group was instructed to play through an adapted version of the educational game Monkey Tales in three weeks' time (gaming group). A second group was instructed to complete a set of math drill exercises in the same period (paper exercises group).

Stimulus material

Educational Game

We used the 3D video game Monkey Tales (Larian studios, 2011), which exists in different versions for second to sixth grade and is used to support the learning of math. The main goal of this educational game is not to instruct but to improve mental arithmetic skills of children by motivating them to engage in drill exercises with increasing time pressure. Only by being faster than a monkey (artificial intelligence) they can go through all the game levels. 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 is meant for children in the 3rd grade (ages 8+) to repeat what they have learned in the 2nd.

The educational game is divided into chapters and levels in which the player has to solve 3D puzzles (moving something that blocks the way or neutralize a laser for instance) and is challenged by a monkey to take part in a minigame (an educational math exercise in classic game format, e.g. 2D shoot 'em up) which the player has to win to get to the next level (see figure 1).

For the present research we made two important modifications on the balancing algorithm of the commercial version in order to make the game as comparable as possible with the paper exercises. First of all, the game was adapted in a way that the player didn't needed to win against the monkey to get to the next level. This was a crucial manipulation to ensure that the quantity of exercises that children received was exactly the same in the two conditions (game vs. paper exercises) since each time that children need to replay a minigame they are performing more exercises. Secondly, we fixed the order of presentation and the type of minigames for each of the 42 basic levels and the final level of the game. The order was established based on in-game logs of five children (second graders without clinically diagnosed learning disorders) who played the commercial version of the game the year before during the same period (data collected for the research reported on Nuñez Castellar et al. 2013). This in order to ensure that the order and difficulty of the exercises and the frequency of the minigames resembles the one that children would get when using the commercial version of the game.

Insert Figure 1 about here

Paper exercises

As one of the goals of the present study was to compare playing Monkey Tales with paper exercises, the latter needed to be as similar as possible to the former. In fact the unique difference between the paper exercises condition and the game condition, should be the game elements of the later. Therefore, the same predetermined order of exercises used in the game, was embedded for the paper exercises. In order to keep the same type of questions, the format of presentation of arithmetic problems in the minigames was respected and accordingly used in the paper exercises (e.g minigames with multiple choice questions were translated into multiple choice exercises).

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 post-test.

Measures

Math performance: accuracy and speed

Two equivalent versions of exams (test A and test B) based on the academic curriculum for second grade in Belgium for assessing the math skills of children were used in the present research (both versions can be found in Nunez Castellar et al. 2013). We used the questions of these two tests to program a computerized version to automatically measure not only the accuracy rates but also the reaction times of each item of the test in milliseconds. We programmed this computerized version using Tscope that is a C/C++ experiment programming library for cognitive scientists (Stevens, Lammertyn, Verbruggen, & Vandierendonck, 2006). After a number of practice trials to become familiar with the multiple choice task, all children performed the computer math test in the pre- and post-session. In each group, half of the children performed test A as pre-test measurement and test B as post-test measurement. The other half performed the tests in the opposite order.

Working memory

Standardized assessments of children's working memory were conducted in the pre and post-test session by means of the Digit Span subtest of the WISC-III NL (Kort et al., 2002), which is an IQ test for children aged 6-17 years standardized for Dutch and Flemish population.This subtest requires a child to repeat 15 series of increasingly long sequences of numbers that are spoken live-voice by an experimenter- at a rate of approximately one digit per second- either forward (8 series) or backward (7 series).

Planning and visuo-motor skills

Standardized assessments of children's planning and visual-spatial skills were performed by means of the Mazes subtest of the WISC-III NL (Kort et al., 2002),. This test consists of 10 mazes of increasing difficulty. Children are required to find their way out of the mazes, this in a limited time of 30, 45 , 60 or 150 seconds.

Enjoyment

Self-reported enjoyment

In the second session, 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 were taken from the extended short feedback questionnaire (Moser et al., 2012) . 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.

Relative enjoyment scale

The Relative enjoyment scale (RES: Jan van Looy et al., manuscript in preparation*)* was used to quantify enjoyment. The RES scores were used for the present research given that when compared with scores of likert-type items included in scales like the “Smileyometer” (Fun Toolkit: Read & MacFarlane, 2000) or the “Funometer” (Extended short feedback questionnaire: Moser eta al., 2012), the RES scores have the advantage to provide scores normally distributed and it is less sensitive to the effects of social desirability. This was important for the present study given that we sought to conduct correlation analyses using the variable enjoyment as continuous measure. Briefly, to complete the RES children are requested to compare the enjoyment of a target activity (e.g. playing Monkey Tales) with another activity (e.g. going to the beach) (see figure 1). During the test the examiner shows sequentially the drawings of two activities and ask: “What do you enjoy more? Playing monkey tales or going to the beach?”. In the middle of the activities a curved scale with seven strips can be observed. Children are instructed to point with the finger in the middle of the curve if both activities are equally nice but otherwise move the finger gradually towards the activity that they like the most. The scale has 12 items and each item can be scored between 1-7 (7 = finger points to the closest strip to the target activity). The 12 activities has been selected of a pool of 25 activities that have been rated in terms of enjoyment by second graders (See figure 1 left). The side where the target activity is presented was counterbalanced (half of the times appear at the right side and vice-versa) to prevent that the results are contaminated by any kind of line bisection bias. Also a training phase was included in which children get familiarized with the scale. Importantly, the results of the RES have been shown to have accurate internal and ecological validity (Jan van Looy et al., manuscript in preparation*)*.

Insert Figure 2 about here

Procedure

The participants were tested at the beginning of May 2013 for the pretest session and at end of May/beginning of June for the posttest session. All pretest and posttest sessions were carried out on the University campus [name left out for review integrity]. We used three rooms for testing: one for parents, one where children performed the computer math test and one for the cognitive and enjoyment tests. Accordingly, three researchers were assigned to each room to supervise the tests.

As previously described, children were randomly assigned to two groups. One group of children was instructed to finish the adapted version 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.

During the same three weeks' time, a second group of children was instructed to complete a set of math drill exercises (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.

All parents received the instruction to let children continue to do their math homework as usual but to not no allow children to play any other kind of digital game during this period. The group that completed the paper exercises received the educational game at the end of the posttest as a reward. All the parents received 15 euros for their participation.

Data Analyses

We calculated descriptive statistics of the socio-demographic variables of each group and controlled that the random allocation of individuals had not involuntarily resulted in systematically biased groups. Having controlled for this, the accuracy rates of the math test, time-to completion of the math test, scaled scores of the Digit span and the scaled scores of the Mazes were subjected to a 2 x 2 repeated measures ANOVA. The variable Group (Gaming vs. Paper exercises) was used as a between-subjects factor and Session (pretest vs. posttest) as a within-subjects factor. Subsequently, post-hoc tests were performed based on Tukey's HSD (Honestly Significant Difference), to correct for multiple comparisons.

To analyse the enjoyment measures two different analyses were conducted. First in order to assess whether significant differences can be observed in the frequency to which children select positive and negative attributes to describe their experience of playing the educational game or solving the paper exercises a chi-square test was performed. Secondly, to investigate differences between groups in the RES scores we conducted a one way ANOVA. For all parametric and non-parametric tests, a significance level of 0.05 was used.

Results

Accuracy math test

A repeated measures ANOVA revealed a main effect of Session, with higher accuracy percentages in posttest when compared with the pretest session ($F(1,50) = 6,74$, $p < 0.05$). The main effect of Group was not significant ($F(1,50) = 0,04$, $p = 0.84$). More importantly the results showed a close to significant interaction between Session and Group ($F(1,50) = 2,86$, $p = 0.09$). Post-hoc comparison showed significant differences between the pretest and the posttest measures only for the group that played the educational game ($p < 0.05$). This indicates 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 (see figure 3).

Insert Figure 3 about here

Time to completion math test

The time-to-completion was subject to a repeated measures ANOVA. Our results indicated a significant main effect of the variable Session, showing a significant reduction of the time-to completion in the posttest measure ($F(1,50) = 44,33$, $p < 0.001$). The main effect of Session, $F(1,50) = 0.21$, $p = .64$ and the interaction between Group and Session, $F(1, 50) = 0.29$, $p = .65$, were not significant. However Post hoc comparisons indicated that both the group that played the educational game and the paper exercises group were significantly faster in the post test session ($p < 0.001$). This suggest that although both groups performed significantly faster in the post-test session the group of children that played the educational game did not outperformed the paper group in terms of speed of completion (see figure 4).

Insert Figure 4 about here

Digit Span

The scaled scores of the digit span scores were also subjected to a repeated measures ANOVA. The results of this analysis revealed non-significant main effects of Session ($F(1,50) = 0,45$, $p = 0.50$) or Group ($F(1,50) = 0,16$, $p = 0.69$). The interaction between Session and Group was also not significant ($F(1,50) = 1,36$, $p = 0.25$).

However a more detailed analyses of the changes in the distribution of the scores between the pre-test and post-test session revealed some differences between the groups when considering individual cases scoring one and a half standard deviation above the mean. Figure 5B shows distribution of the pre- vs posttest scaled scores for the paper exercises and the gaming group. The normal distributions for the pre and posttest scores are shown as grey and red bell shaped curves. The number of individual cases per score are also plotted. As it can be observed in the figure on the gaming group before training 4% of the subjects scored above 14 while in the second session 19% did it (15% more than in the pretest). On the contrary on paper group only moderate gains are observed. Before training 16 % of the subjects scored above 24% while in the second session 32% (8% more than in the pretest). In other words these changes in distribution indicates the percentage of participants who achieved outstanding scores in the digit span subtest was twice more after the game training when compared with the traditional training.

Insert Figure 5 about here

Mazes

The scaled scores of the scores of the mazes subtest were also subjected to a repeated measures ANOVA. The results of this analysis revealed non-significant main effects of Session ($F(1,50) = 0,22$, $p = 0.64$) or Group ($F(1,50) = 0,01$, $p = 0.92$). The interaction between Session and Group was also not significant ($F(1,50) = 0,67$, $p = 0.41$).

A more detailed analyses of the changes in the distribution of the scores between the pre-test and post-test session did not revealed some differences between the groups when considering individual cases scoring one and a half standard deviation above the mean. Figure 6B shows distribution of the pre- vs posttest scaled scores for the paper exercises and the gaming group. The normal distributions for the pre and posttest scores are shown as grey and red bell shaped curves. The number of individual cases per score are also plotted. As it can be observed in the figure on the gaming group before training 15% of the subjects scored above 22 and equally in the second session 15% did it (no gains). On the paper exercises group before training 8% of the subjects scored above 22 while in the second session 0% did it (no gains). This suggests that when considering the number of subjects who reached outstanding scores no differences are observed between the gaming and the paper exercises group.

Insert Figure 6 about here

Enjoyment

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 Figure 7. The results of a chi-square test revealed that the frequency at which the attributes 'exciting', 'boring' 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 = 52) = 5.97, $p < .05$. Likewise, more children selected the attribute "exciting" to describe their experience playing the educational game as compared to solving the paper math drills, χ^2 (1, N = 52) = 4.54, $p < .05$. Finally the results also revealed that children more frequently reported the attributes 'simple', χ^2 (1, N = 52) = 7.86, $p < .01$, 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.

Insert Figure 7 about here

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 (see Figure 8). The results of a chi-square test revealed that the proportion of participants who stated they wouldn't like to repeat the experience tended to be higher in the group that did the paper math drills (24%) than in the group that played the educational game (7%), χ^2 (1, N = 52) = 2.75, $p = .09$ (see Figure 6). Accordingly the proportion that answered yes and maybe to the question whether they would like to repeat the experience again was larger in the gaming group (93%) than in the paper exercises group (76%).

Insert Figure 8 about here

Correlations

We conducted correlation analyses among all the continuous measurements described above. The results can be observed in the Table 2. Our results revealed only a significant correlation between improvements in working memory scores and reaction times on the math test. The negative correlation indicates that the larger the improvement on the digit span scores the faster that participants completed the math test in the post test session.

Insert Table 2 about here

Discussion

In this follow-up study we explore whether traditional methods and game training differ in terms of the cognitive processes that both are able to impact. We incorporated standardized measures of working memory and visuo-motor skills, and applied different measures of affective and cognitive learning outcomes. Our results point to some interesting findings. First, in terms of affective outcomes our results showed that an important difference between traditional methods and game training is that the later elicits more positive affective responses. The results of the relative enjoyment scale showed that children reported that playing the game was more enjoyable than many other activities when compared with doing paper exercises. Likewise, when participants 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 “exciting” while a significantly higher proportion of the children who filled out the paper exercises described their experience as “boring” and “easy”. Additionally, 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 the findings our previous study assessing the effectiveness of the commercial game Monkey Tales (Nunez Castellar et al., 2013) and with other studies showing that children who play games aimed to train arithmetic performance report more enjoyment than the children who use traditional methods (Ke, 2008; Koran & McLaughlin, 1990).

Secondly, another interesting issue explored in this study was whether the enjoyment scores of the relative enjoyment scale (affective measure) were related with the gains in cognitive outcomes. According to the framework of Kraiger, Ford and Salas (Kraiger, Ford, & Salas, 1993) learning outcomes are not discrete but are usually interacting; changes in cognitive outcomes could for instance co-occur with changes in affective outcomes. Here we investigated whether besides co-occurring, the enjoyment scores had a predictive value for the cognitive learning gains. We failed to find evidence supporting this idea. Our results showed that enjoyment scores were not significantly correlated with the gains in the cognitive measures included in the present study. Although this is a finding that should be interpreted with caution, considering that the RES is a recently developed scale to measure enjoyment in children, we think that this remains an interesting venue for future research. An unresolved issue in the field is that, as indicated by Lumby (2011), enjoyment could be conceived as a precursor, a parallel experience, a result of learning or all three (Lumby, 2011). Future studies aimed to probe the potential of serious games as learning tools, should provide evidence about how enjoyment can improve learning over and above the training effects expected from the classical methods.

Thirdly, a novel element in this study was the inclusion of standardized assessments of children's working memory and visuo-motor skills before and after game training when compared with traditional training. An important finding of the present study is that our results showed a significant correlation between the gains in the working memory scores and the changes in the time required to complete a computer math test. The larger the working memory improvement, the faster that children completed the test in the post test session.

This finding is in line with previous studies showing that working memory capacity closely relates to skill in arithmetic's and, in particular, to the speed to solve arithmetic problems (Geary & Widaman, 1992; Lemaire et al., 1996; Rubinsten & Henik, 2009).However, an issue that deserves attention is that the gains in the scores of the Digit span test were achieved through methods not aimed at formal working memory training for children. In the last decade working memory trainings have become popular as tools for improving cognitive ability and scholastic attainment, particularly in developing children and adults (for a complete review see Melby-Levarg & Hulme, 2013). Therefore establishing whether games as educational tools can also impact working memory skills is an issue that should be further investigated. A final consideration in relation with this finding is that this effect was only found for the Digit span measure and not for the scores in the subtest of Mazes, which indicates a dissociation between the possibilities that games offer to train visuo-motor skills and working memory. In fact we did not find any evidence of gains in the scores of the mazes when comparing the pretest and the posttest performance for both groups.

At this point, we would have to speculate about the precise relationship between game training and working memory capacity and how it differs from traditional methods. First of all, we would like to point out that the present study can provide only an indication that children might be able to benefit from games for improving working memory skills more than from classical methods. Specifically, we found that when examining changes in the distribution of the working memory scores between the pre and the post-test sessions, the percentage of participants who achieved outstanding scores was twice as much after the game training than after the traditional training. However the results of the repeated measures ANOVA did not reveal significant differences between groups. Further research with larger sample sizes should be conducted to establish whether we failed to find significant effects when considering the averages due to the small sample size of this study. Also, although in the present study only one measure of working memory capacities was included, future research could conduct a more detailed evaluation of working memory improvements

Finally, in relation to the analyses of the accuracy and time-to-completion improvements, the results of the present study generally replicate the results reported in Nuñez Castellar et al., 2013, using the same math test and the commercial version of Monkey Tales. Only the accuracy improvement percentage of the previous study was slightly larger for both groups when compared with improvements reported in the present one. This can be due to the fact that both the game and the paper exercises were slightly modified for the present study. We adapted the material to ensure that only one difference was present between conditions; the game elements. As described before each single math exercise, type of question (e.g., multiple choice) and the order of presentation was matched. Although this manipulation is an important improvement for the assessment of game training versus traditional training, it has the drawback that it makes difficult a direct comparison between the results of both studies.

To conclude the present work provided some preliminary evidence suggesting that the underlying mechanisms of performance gains might be different between traditional learning methods and game training. Consequently, several questions could now be experimentally investigated by future research, for instance the extent to which results are generalizable to other game, whether similar positive effects can be found with children of different ages, and whether high and low high achievers can benefit to the same extent from game training.

6. Acknowledgements

This research was conducted in the context of the iMinds-ICON 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) for their support with the materials and information about the educational game.

References

American Psychiatric, A. & American Psychiatric Association. Task Force On, D.-I., 2000. *Diagnostic and statistical manual of mental disorders : DSM-IV-TR* Washington, DC: American Psychiatric Association.

Baddeley, A., 2000. The episodic buffer: a new component of working memory? *Trends in Cognitive Sciences,* 4**,** 417-423.

Bavelier, D., Green, C.S., Pouget, A. & Schrater, P., 2012. Brain Plasticity Through the Life Span: Learning to Learn and Action Video Games. *In* S.E. Hyman (ed.) *Annual Review of Neuroscience, Vol 35.* 391-416.

Boot, W.R., Kramer, A.F., Simons, D.J., Fabiani, M. & Gratton, G., 2008. The effects of video game playing on attention, memory, and executive control. *Acta Psychologica,* 129**,** 387-398.

Campbell, T., Dollaghan, C., Needleman, H. & Janosky, J., 1997. Reducing bias in language assessment: Processing-dependent measures. *Journal of Speech Language and Hearing Research,* 40**,** 519-525.

Colzato, L.S., Van Den Wildenberg, W.P., Zmigrod, S. & Hommel, B., 2012. Action video gaming and cognitive control: playing first person shooter games is associated with improvement in working memory but not action inhibition. *Psychological research***,** 1-6.

Diamond, A. & Lee, K., 2011. Interventions shown to aid executive function development in children 4 to 12 years old (vol 333, pg 959, 2011). *Science,* 334**,** 311-311.

Elson, M., & Bente, G, 2009. *CORTEX - Computer-Aided Registration Tool for Experiments* [online]. University of Cologne. Available from: http://cortex.uni-koeln.de [Accessed Access Date

Geary, D.C. & Widaman, K.F., 1992. Numerical cognition - on the convergence of componential and psychometric models. *Intelligence,* 16**,** 47-80.

Gredler, G.R., 1990. Sattler, J. M. (1988). Assessment of children (3rd ed.). San Diego: Jerome M. Sattler, Publisher, 995 pp. . *Psychology in the Schools,* 27**,** 276-277.

Green, C.S. & Bavelier, D., 2003. Action video game modifies visual selective attention. *Nature,* 423**,** 534-537.

Holmes, J., Gathercole, S.E. & Dunning, D.L., 2009. Adaptive training leads to sustained enhancement of poor working memory in children. *Developmental Science,* 12**,** F9-F15.

Kail, R. & Park, Y.S., 1990. Impact of practice on speed of mental rotation. *Journal of Experimental Child Psychology,* 49**,** 227-244.

Karle, J.W., Watter, S. & Shedden, J.M., 2010. Task switching in video game players: Benefits of selective attention but not resistance to proactive interference. *Acta Psychologica,* 134**,** 70-78.

Ke, F.F., 2008. A case study of computer gaming for math: Engaged learning from gameplay? *Computers & Education,* 51**,** 1609-1620.

Klingberg, T., Fernell, E., Olesen, P.J., Johnson, M., Gustafsson, P., Dahlstrom, K., Gillberg, C.G., Forssberg, H. & Westerberg, H., 2005. Computerized training of working memory in children with ADHD - A randomized, controlled trial. *Journal of the American Academy of Child and Adolescent Psychiatry,* 44**,** 177-186.

Koran, L.J. & Mclaughlin, T.F., 1990. Games or drill: Increasing the multiplication skills of students. *Journal of Instructional Psychology* 17**,** 222-230.

Kort, D. W., Schittekatte, M., Compaan, E. L., Bosmans, M., Bleichrodt, N., Vermeir, G., ... & Verhaeghe III, P. (2002). WISC-III NL. Handleiding. London: The Psychological Corporation.

Kraiger, K., Ford, J.K. & Salas, E., 1993. Application of cognitive, skill-based, and affective theories of learning outcomes to new methods of training evaluation. *Journal of Applied Psychology,* 78**,** 311-328.

Kurdek, L.A. & Sinclair, R.J., 2001. Predicting reading and mathematics achievement in fourth-grade children from kindergarten readiness scores. *Journal of Educational Psychology,* 93**,** 451-455.

Lemaire, P., Abdi, H. & Fayol, M., 1996. The role of working memory resources in simple cognitive arithmetic. *European Journal of Cognitive Psychology,* 8**,** 73-103.
Lumby, J., 2011. Enjoyment and learning: policy and secondary school learners' experience in England. *British Educational Research Journal,* 37**,** 247-264.
Mclean, J.F. & Hitch, G.J., 1999. Working memory impairments in children with specific arithmetic learning difficulties. *Journal of Experimental Child Psychology,* 74**,** 240-260.
Moser, C., Fuchsberger, V., & Tscheligi, M. (2012). *Rapid assessment of game experiences in public settings*. Paper presented at the Proceedings of the 4th International Conference on Fun and Games, Toulouse, France.
Nouchi, R., Taki, Y., Takeuchi, H., Hashizume, H., Nozawa, T., Kambara, T., Sekiguchi, A., Miyauchi, C.M., Kotozaki, Y., Nouchi, H. & Kawashima, R., 2013. Brain Training Game Boosts Executive Functions, Working Memory and Processing Speed in the Young Adults: A Randomized Controlled Trial. *PLoS ONE,* 8**,** e55518.
Núñez Castellar, E., Van Looy, J., Szmalec, A. & De Marez, L., Improving arithmetic skills through gameplay: Assessment of the effectiveness of an educational game in terms of cognitive and affective learning outcomes. *Information Sciences*. In press
Pisoni, D.B., Kronenberger, W.G., Roman, A.S. & Geers, A.E., 2011. Measures of digit span and verbal rehearsal speed in deaf children after more than 10 years of cochlear implantation. *Ear Hear,* 32**,** 60S-74S.
Read JC, MacFarlane SJ (2000) Measuring fun—usability testing for children. Computers and Fun 3, *BCS HCI Group*, York, England
Rubinsten, O. & Henik, A., 2009. Developmental Dyscalculia: heterogeneity might not mean different mechanisms. *Trends in Cognitive Sciences,* 13**,** 92-99.
Siegel, L.S. & Feldman, W., 1983. Nondyslexic children with combined writing and arithmetic learning-disabilities. *Clinical Pediatrics,* 22**,** 241-244.
Skuse, D. H. (Ed.). (2003*). Child Psychology and Psychiatry: An Introduction*. The Medicine Publishing Company.
Sortor, J.M. & Kulp, M.T., 2003. Are the results of the Beery-Buktenica Developmental Test of Visual-Motor Integration and its subtests related to achievement test scores? *Optometry and Vision Science,* 80**,** 758-763.
Stevens, M., Lammertyn, J., Verbruggen, F. & Vandierendonck, A., 2006. Tscope: AC library for programming cognitive experiments on the MS Windows platform. *Behavior Research Methods,* 38**,** 280-286.
Studios, L., 2011. Monkey tales. Gent, Belgium: Die keure & Larian Studios
Turley-Ames, K.J. & Whitfield, M.M., 2003. Strategy training and working memory task performance. *Journal of Memory and Language,* 49**,** 446-468.
Van Looy, J., Houttekier ,E., Núñez Castellar, E. Development and validation of a relative enjoyment scale to measure children's enjoyment. *Manuscript in preparation*

Figure captions

Figure 1. Screen shoot Monkey Tales

Figure 2. Left side: Picture of the pilot study in which 25 activities were rated in terms of enjoyment. Right side: Format of the Relative Enjoyment Scale.

Figure 3. Mean accuracy rates in the pre and post-test measurements by group

Figure 4. Mean time-to-completion in the pre and post-test measurements by group

Figure 5. (A) Mean Digit Span scores in the pre and post-test measurements by group (B) Distribution of the pre- vs post-test scores by group. The normal distributions for the pre and post-test scores are shown as grey and red bell shaped curves

Figure 6. (A) Mean mazes scores in the pre and post-test measurements by group (B) Distribution of the pre- vs post-test scores by group. The normal distributions for the pre and post-test scores are shown as grey and red bell shaped curves

Figure 7. Attributes used by children to describe their experience of playing the educational game or solving the paper exercises

Figure 8 (A) Mean scores Relative Enjoyment Scale (B) Proportion of children who would like to play the educational game or do more math exercises again

	Educational Game (N=27)	Paper Exercises (N=25)		
	n	**n**	**Chi²**	**p**
Male gender	19	17	0.03	.85
	Mean	**Mean**	**t**	**p**
Age	7.40	7.44	-0.21	.82
Level education parents	**Median**	**Median**	***U***	**p****
Education level father*	4	4	328	.87
Education level mother*	4	4	334	.95
Study and game habits	**Mean**	**Mean**	**t**	**p**
Homework hours per week	2h 38min	2h 02min	1.27	.21
Math homework hours per week	1h 12min	1h 08min	0.27	.78
Gaming hours during the week	3h 48min	2h 50min	0.83	.40
Gaming hours during the weekend	2h 33min	1h 48min	1.67	.10

*Four levels: Primary = 1, Junior High School/Middle School = 2, High School = 3, College/University = 4.

** Independent Sample Mann-Whitney U Test

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

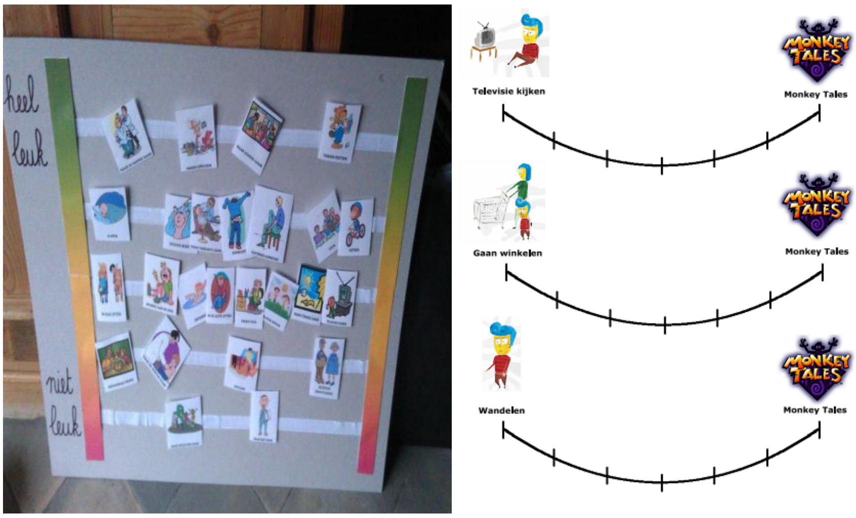
	Accuracy	Reaction Times	Digit Span	Mazes	RES score
Accuracy	1,0000	,2261	,0336	-,1732	,2052
Reaction times		1,0000	-,3104**	,0182	-,0224
Digit Span			1,0000	-,2165	,0135
Mazes				1,0000	-,0177
RES score					1,0000

*p < .05. **p < .001

Table 2. Correlation analyses among all cognitive and affective learning outcomes.

Figure 1

Figure 2



heel
leuk
niet
leuk
Televisie kijken
Monkey Tales
Gaan winkelen
Monkey Tales
Wandelen
Monkey Tales

Figure 3

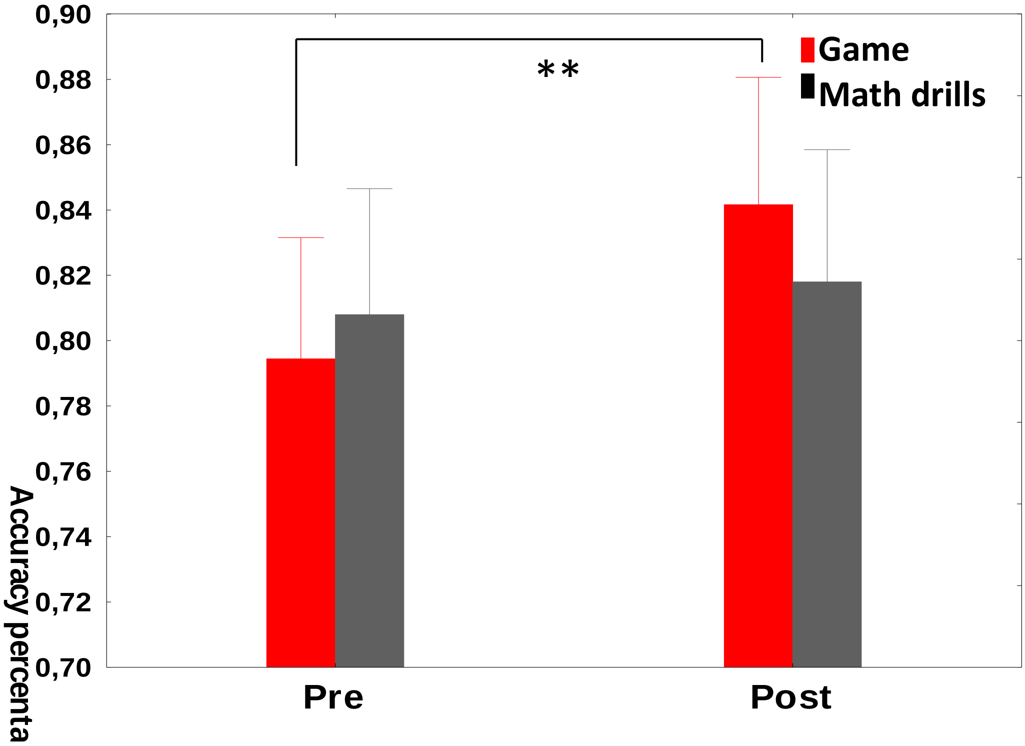

0,90
0,88
0,86
0,84
0,82
0,80
0,78
0,76
0,74
0,72
0,70
Accuracy percenta
**
Game
Math drills
Pre
Post

Figure 4

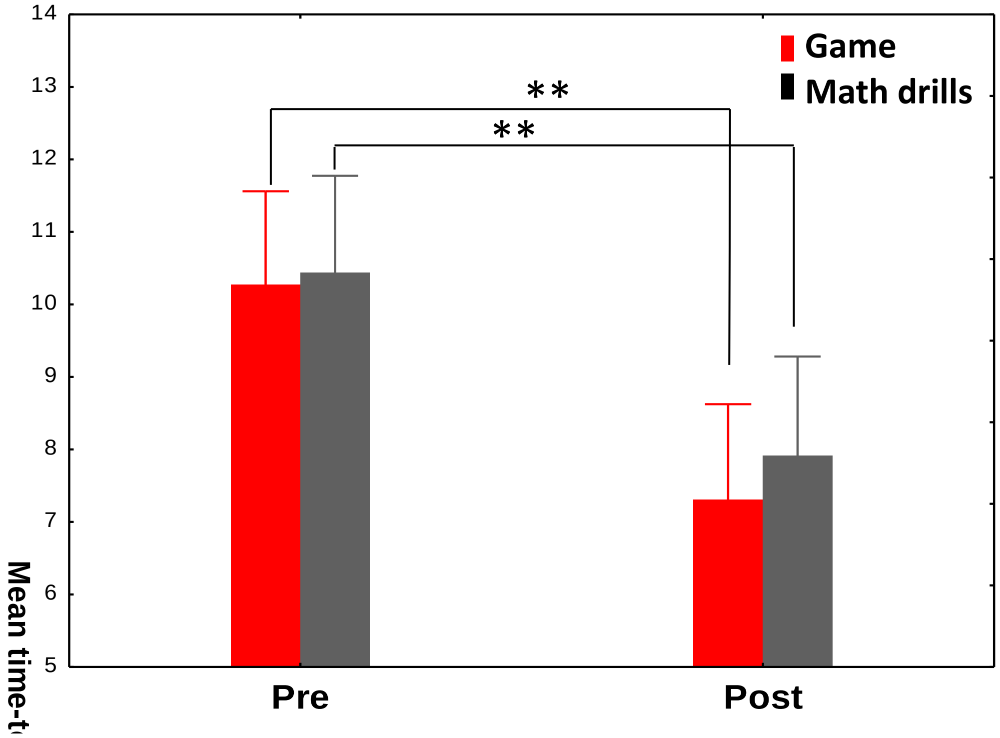
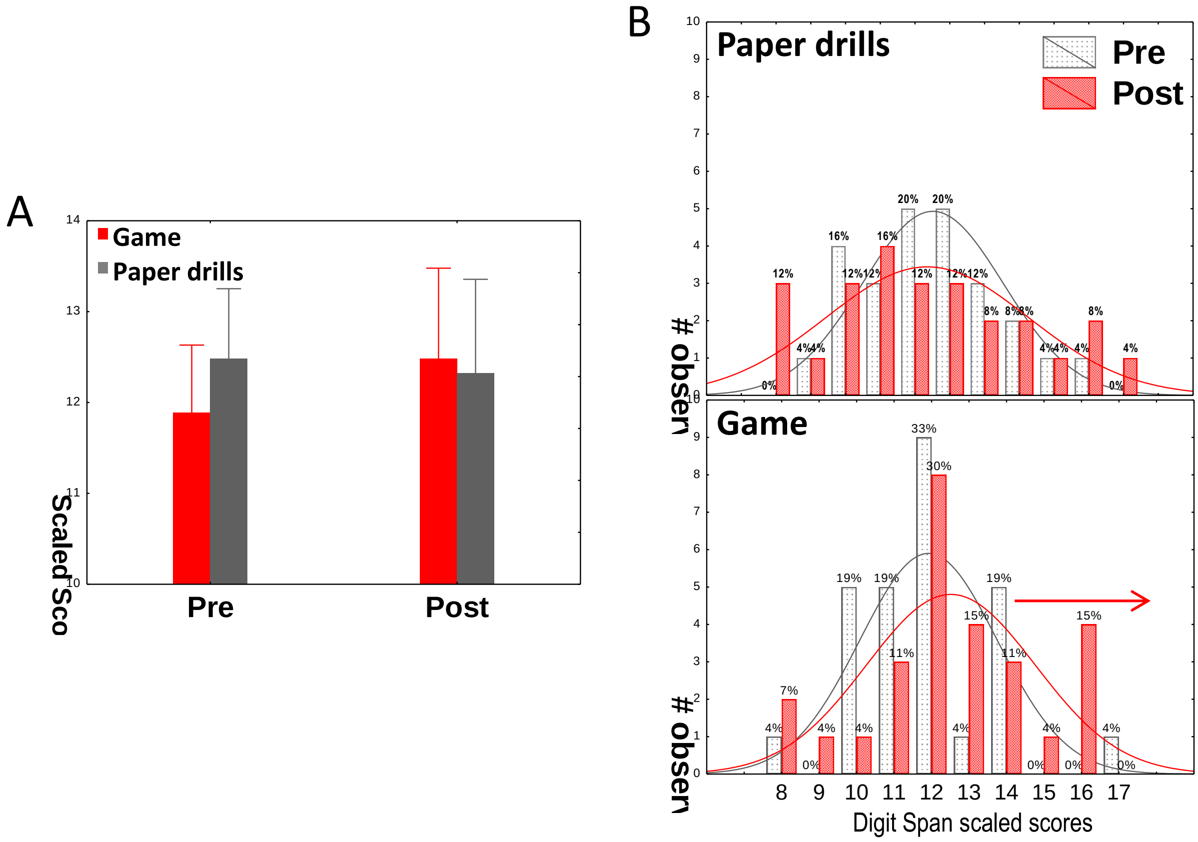
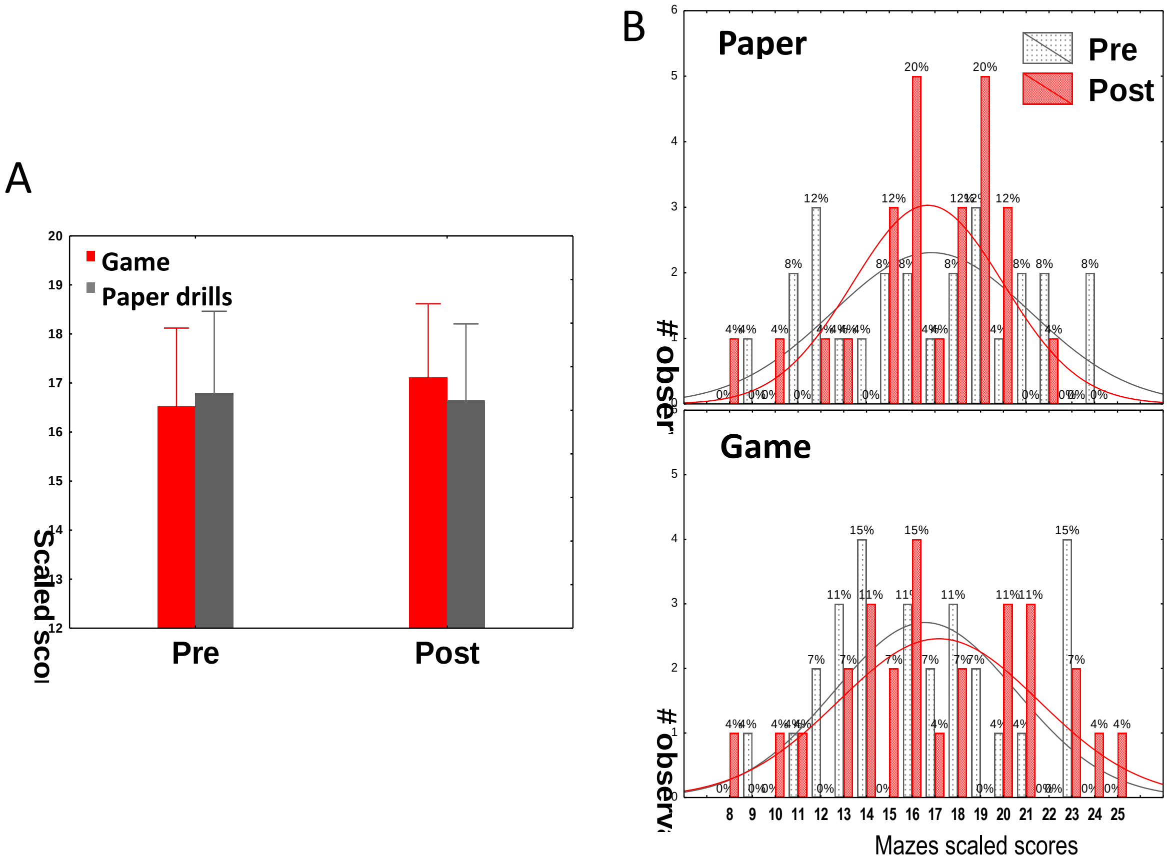


Figure 5



A
Game
Paper drills
Scaled Sco
Pre
Post
B
Paper drills
Pre
Post
Game
obser
Digit Span scaled scores

Figure 6



A
Game
Paper drills
Scaled scor
Pre
Post
B
Paper
Pre
Post
Game
obser
Mazes scaled scores

Figure 7

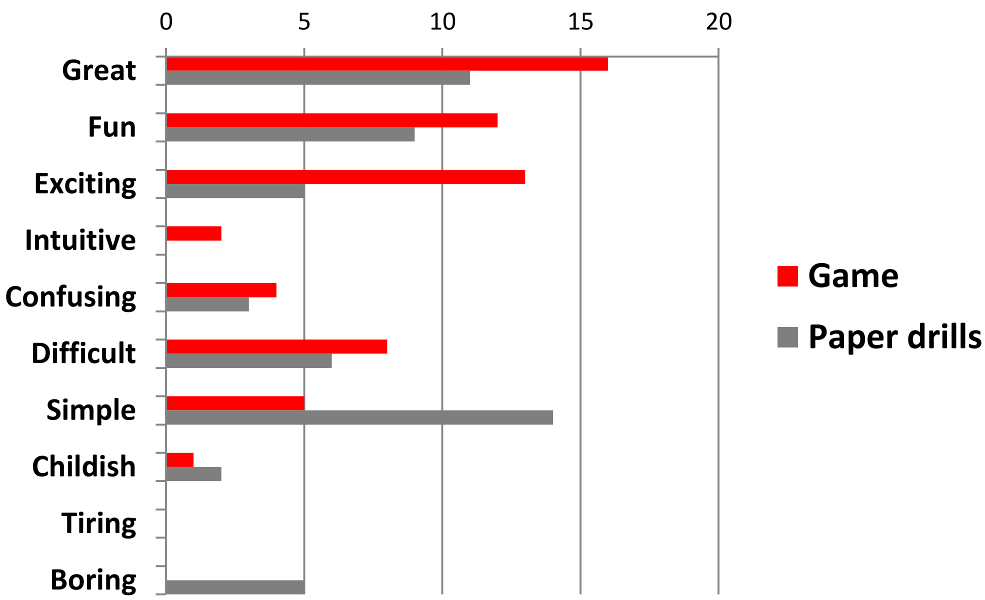
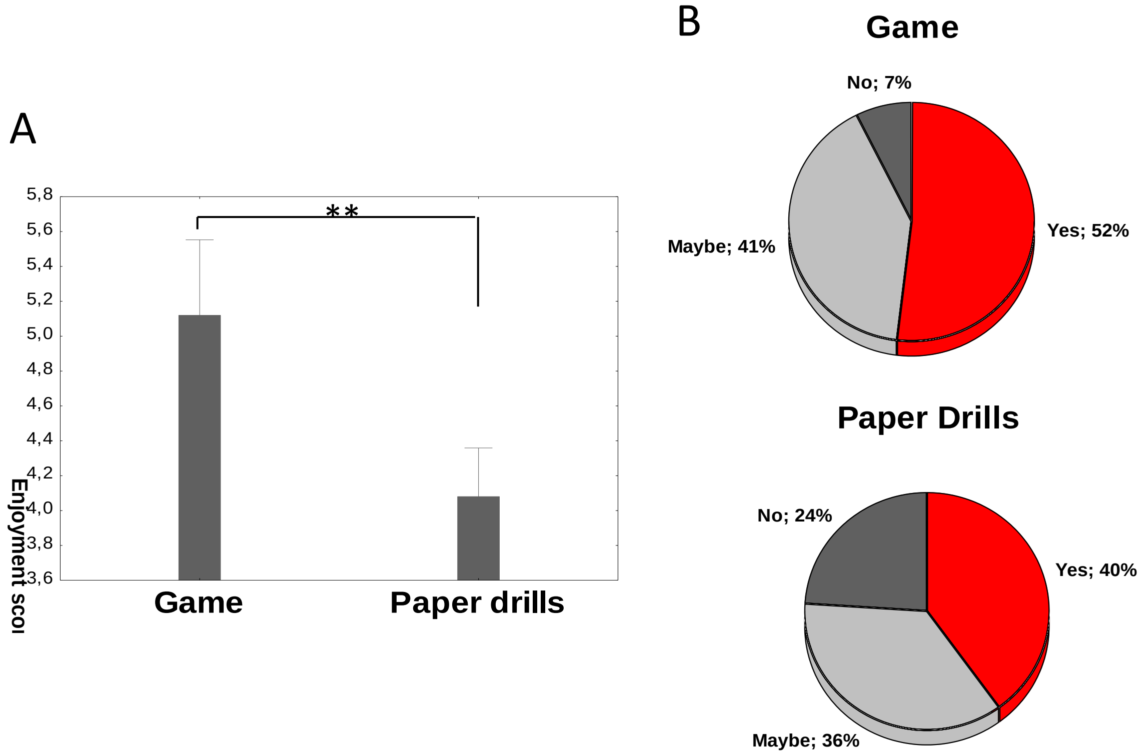

0
5
10
15
20
Great
Fun
Exciting
Intuitive
Confusing
Difficult
Simple
Childish
Tiring
Boring
Game
Paper drills

Figure 8


A
5,8
5,6
5,4
5,2
5,0
4,8
4,6
4,4
4,2
4,0
3,8
3,6
Enjoyment scoi
**
Game
Paper drills
B
Game
No; 7%
Maybe; 41%
Yes; 52%
Paper Drills
No; 24%
Yes; 40%
Maybe; 36%