Introduction

An important advantage of serious games over traditional educational tools, is that in the former customized content can be provided to learners based on adaptive algorithms. Research into the effects of adaptive content is sparse however. Klinkenberg, Straatemeier & van der Maas (2010) collected data of 3648 children, who completed about 33% of arithmetic problems outside school hours with the game platform Math Garden and showed that the adaptive system offered better measurement precision, high validity and reliability and high pupil satisfaction. Interestingly in a follow-up study the relationship between success rates (60%, 75%, and 90%: manipulated by means of the adaptive algorithm), math anxiety and perceived math competence was experimentally investigated (Jansen et al., 2013). The results showed that improvements in math performance were most pronounced for the condition with the highest success rate but that this manipulation did not have any effect on math anxiety or perceived math competence.

The present research aimed to assess the impact of the rule-based adaptive system of the commercial version of the game Monkey tales on learning outcomes. We investigated whether objective measures of cognitive (math performance, working memory & visuo-motor skills) and affective (enjoyment) learning outcomes, could provide evidence of the effectiveness of the algorithm for two different user profiles: learners with high arithmetic skills (high achievers) versus learners with low arithmetic skills (low achievers).

Method

Participants

In the first evaluation, 66 second graders (45 boys and 21 girls) were tested. At the second evaluation (posttest), only 62 children could be assessed. From this sample, participants clinically diagnosed learning disability, ADHD, and dyslexia, and who performed the computer math test at chance level or below were excluded. The data reported here include 53 children.

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 in which the adaptive algorithm was disabled (unbalanced condition). A second group was instructed to play through the commercial version of Monkey tales - with the adaptive algorithm- in the same period (balanced condition).

Stimulus material

Educational game

We used the 3D video game *Monkey Tales: Museum of Anything* (Larian studios, 2011, which is meant for children in the 3rd grade to repeat what they have learned in the 2nd. The educational game is divided into 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 2D minigame (an educational math exercise).

Unbalanced condition

For this condition we disabled the adaptive algorithm and 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 to ensure that the order and difficulty of the exercises and the frequency of the minigames is similar to the commercial version.

Measures

Kortrijkse Arithmetic Test Revised 2006 (KRT-R)

This is a standardised mathematics test on domain-specific knowledge and skills, resulting in a percentile score on mental calculation and system knowledge (Kortrijkse Rekentest, KRT; Cracco et al. 1995). Based on these scores two user profiles were established. We conducted a median split to identify high and low achievers.

Math performance: accuracy and speed

Two equivalent computerized 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 (both versions can be found on [Reference removed to protect anonymity]).

Working memory and Planning and visuo-motor skills

Standardized assessments of children's working memory and visual-spatial skills were conducted in the pre- and post-test session by means of the Digit Span and the Mazes 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.

Relative enjoyment scale

The Relative enjoyment scale (RES: Jan van Looy et al., manuscript in preparation) was used to quantify enjoyment. The RES scores have the advantage of providing normally distributed scores and of being less sensitive to social desirability effects. Importantly, the results of the RES have been shown to have accurate internal and ecological validity (see figure 1).

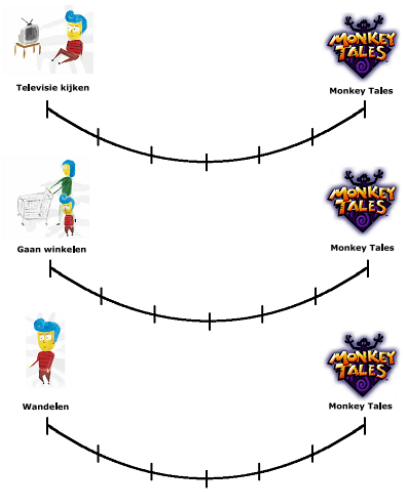

Figure 1. Fragment of the Relative Enjoyment Scale.

Results

Accuracy math test

The results of the mixed ANOVA revealed a significant main effect of Session, indicating higher accuracy rates in the post test session when compared with the pre-test session ($F(1,48) = 20,34$, $p < 0.001$ and also of User Profile showing increased accuracy in the computer math test for the high compared with the low achievers, which is an expected finding ($F(1,48) = 40,57$, $p < 0.001$). No other significant main effects or significant interactions were found.

Time-to-completion math test

The results of the mixed ANOVA revealed a main effect of Session, indicating faster reaction times in the post test session for both conditions ($F(1,48) = 66,86$, $p < 0.001$). Also a significant main effect of User profile was found showing faster reaction times in the computer math test for the high compared with the low achievers ($F(1,48) = 6,51$, $p = 0.01$). No other significant main effects or significant interactions were found.

Working memory

Our results showed a main effect of Session, indicating that children achieved higher scores in the Digit Span in the post-test session ($F(1,50) = 10,83$, $p < 0.05$). No other significant main effects or significant two way interactions were found but importantly a marginally significant three way interaction was revealed ($F(1,50) = 6,37$, $p = 0.07$). This interaction showed that low achievers had a larger improvement in working memory scores by playing the balanced version of the game (see figure 2).

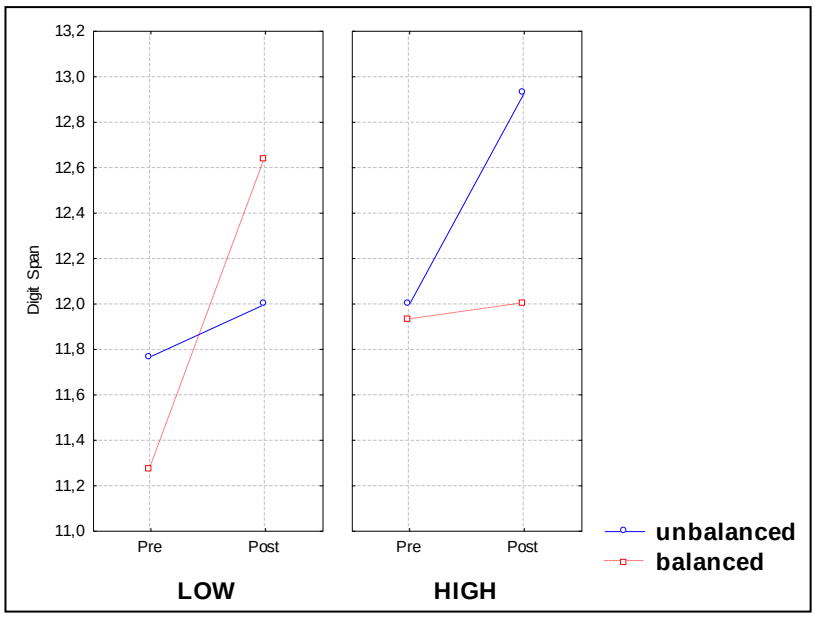

Figure 2. Mean Digit Span scores in the pre and post-test measurements per condition for low and high achievers

Visuo-motor skills

The results of the mixed ANOVA did not reveal any significant main effects or significant interactions.

Enjoyment

The results of the ANOVA did not revealed any significant main effects or significant interaction when considering the RES scores, although by visual inspection it is possible to observe that low achievers who received the unbalanced version reported it as less enjoyable (Figure 3).

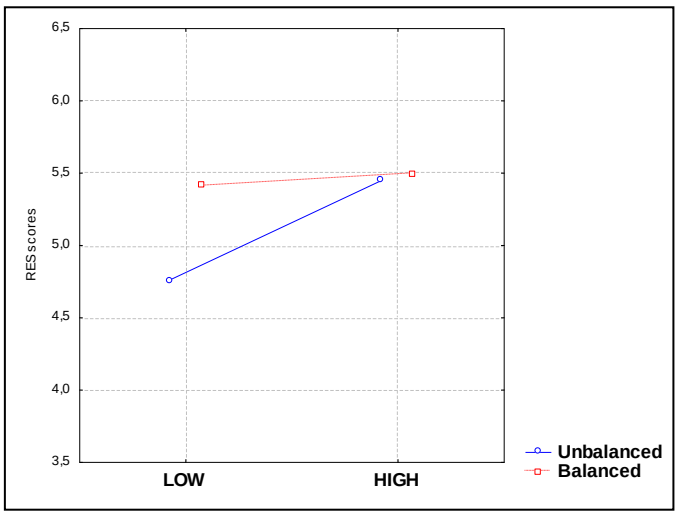

Figure 3. Mean scores Relative Enjoyment Scale

Enjoyment in-game data (logs)

The number of bananas collected by the participants during the game was considered as an enjoyment index and was subject to a factorial ANOVA. The results revealed a main effect of the variable Condition ($F(1,48) = 4,21$, $p < 0.05$), indicating that participants in the balanced condition picked up more bananas than in the unbalanced condition. In other words, they were more motivated to face game challenges when the algorithm was enabled. No other main effects or significant interactions were found.

Math performance in-game data (logs)

In six of the minigames of Monkey Tales children are requested to answer simple arithmetic multiple choice question (e.g. 4 + 6= a.10 b.8 c.12). On average children solved 64 of this this type of questions. The number of errors committed were subject to a factorial ANOVA. The results did not reveal any significant main effects or significant interaction between conditions.

Correlations

We conducted correlation analyses between all the continuous measurements described above. The results can be observed in the Table 1.

	Accuracy	Time-to-completion	Working memory	Mazes	RES	Bananas picked up	In-game math errors
Accuracy	1,0000 p= ---	,3232 p=,019	-,1018 p=,473	-,1144 p=,419	,1784 p=,206	,0704 p=,620	-,0475 p=,738
Time-to-completion		1,0000 p= ---	-,1461 p=,301	-,1282 p=,365	,0933 p=,511	,1113 p=,432	,1115 p=,431
Working memory			1,0000 p= ---	,1360 p=,336	-,2919 p=,036	-,0522 p=,713	-,1539 p=,276
Mazes				1,0000 p= ---	-,1895 p=,178	-,0166 p=,907	,0680 p=,632
RES					1,0000 p= ---	,3323 p=,016	,0801 p=,573
Bananas picked up						1,0000 p= ---	-,1581 p=,263
In-game math errors							1,0000 p= ---

Table 1. Correlation analyses among all cognitive and affective learning outcomes.

Discussion

Our results indicate that low and high achievers show similar improvements in most of the cognitive learning outcomes assessed in the present study. Only working memory improvements were superior for low achievers who played the game with the adaptive algorithm, suggesting that children with low arithmetic skills show a larger benefit of math game training with adaptive content.

Another finding is that on average participants got higher accuracy rates and faster reaction times in the post-test session but equivalent working memory scores. These improvements were not related to user profile (low vs. high achievers) or to the experimental condition (balanced vs. unbalanced). This allows us to conclude that although the adaptive system may have had a positive impact on the learning experience, arithmetical improvements are not dependent on this game feature.

Finally, when considering in-game log data, we found that the number of bananas collected during the game was positively correlated with the scores of the Relative Enjoyment Scale. This information could be potentially useful to identify children who are enjoying the game based on in-game logs. Future research could also for instance investigate to what extent objective measures of enjoyment like in-game logs can be used a predictor of other experience variables and affective outcomes such as attitude towards the subject.

References

Cracco, J., Baudonck, M., Debusschere, A., Dewulf, B., Samyn, F., & Vercaemst, V. (1995). Kortrijkse Rekentest. Kortrijk: Revalidatiecentrum Overleie.

Jansen, B. R., Louwerse, J., Straatemeier, M., Van der Ven, S. H., Klinkenberg, S., & Van der Maas, H. L., 2013. The influence of experiencing success in math on math anxiety, perceived math competence, and math performance. *Learning and Individual Differences* 24, pp. 190-197.

Klinkenberg, S., Straatemeier, M., & van der Maas, H. L. J. (2011). Computer adaptive practice on Maths ability using a new item response model for on the fly ability and difficulty estimation. Computers & Education, 57, 1813-1824.

Kort, W., Schittekatte, M., Dekker, P.H., Verhaeghe, P., Compaan, E.L., Bosmans, M. & Vermeir, G. (2005). WISC-III NL. Handleiding en Verantwoording. London: The Psychological Corporation

Studios, L. (2011). Monkey Tales. Gent, Belgium: Die Keure Larian Studios.

Van Looy, J., Houttekier ,E., Núñez Castellar, E. Development and validation of a relative enjoyment scale to measure children's enjoyment. *Manuscript in preparation*