A Teachable-Agent Arithmetic Game's Effects on Mathematics Understanding, Attitude and Self-efficacy

Lena Pareto[1], Tobias Arvemo[2], Ylva Dahl[3], Magnus Haake[4], and Agneta Gulz[5]

[1] University West, Media & Design, Sweden
lena.pareto@hv.se
[2] University West, Mathematical Statistics, Sweden
Tobias.Arvemo@hv.se
[3] Uddevalla Schools, School Development, Sweden
ylva.dahl@uddevalla.se
[4] Lund University Design Sciences, Sweden
Magnus.haake@design.lth.se
[5] Lund University Cognitive Science, Sweden
agneta.gulz@lucs.lu.se

Abstract. A teachable-agent arithmetic game is presented and evaluated in terms of student performance, attitude and self-efficacy. An experimental pre-post study design was used, enrolling 153 3rd and 5th grade students in Sweden. The playing group showed significantly larger gains in math performance and self-efficacy beliefs, but not in general attitude towards math, compared to control groups. The contributions in relation to previous work include a novel educational game being evaluated, and an emphasis on self-efficacy in the study as a strong predictor of math achievements.

Keywords: teachable agents, mathematics achievement, attitude, self-efficacy.

1 Introduction

Educational games for mathematics have documented effects on learning and motivation [1], [2], [3], [4] and [5]. Games are considered to be effective tools since they are action-based; motivational; accommodate multiple learning styles and skills; reinforce mastery skills; and provide interactive and decision making context [5]. The instructional effectiveness of a game depends both on its particular characteristics and how it is used in classroom instruction [4], [6]. The relation between game characteristics and competence promotion is not well understood [3]. One such characterization is proposed in [7] where the authors claim that technology, such as games, need to be pedagogically sound, mathematically true and cognitively defined in order to deepen understanding. Furthermore, technology should bring reasoning into the environment and allow exploration, conjecture and testing to deepen mathematical understanding [7]. Teachable Agents (TA), i.e., agents that can learn [8], have previously been used to scaffold reflection, conceptual understanding as well as motivation [2], [9]. Below, we will discuss how our Teachable Agent Arithmetic Game (hereafter TAAG) relates to the described characterization.

G. Biswas et al. (Eds.): AIED 2011, LNAI 6738, pp. 247–255, 2011.

In addition to performance, affective issues need to be included in studies of cognition and instruction to have an impact on mathematics education [10]. Attitude, belief, and emotion are the major descriptors of the affective domain, and many mathematics educators consider attitude as their major concern [10]. Attitude toward mathematics refers to: students' affective responses to – their liking or disliking of – mathematics; their tendency to engage in or avoid mathematical activities; their belief in their mathematics ability (i.e., self-efficacy) and their believing that mathematics is useful or useless [11]. Lately, self-efficacy has attracted special attention since self-efficacy beliefs have been shown to be strong predictors of actual accomplishments [12], [13], and [14]. Therefore, we have chosen to study the issues of general attitude towards mathematics and self-efficacy separately.

The primary research questions addressed in this study are: Will TAAG have effects on 1) conceptual understanding of arithmetic, 2) attitudes towards mathematics and 3) self-efficacy beliefs regarding arithmetic performance? As a secondary explanatory question we will explore if achievements are effected by students' self-reported like/dislike of mathematics and/or by different levels of authenticity in the learning situation. To address the latter issue we compare the situation where the game play occur in full class lead by regular teachers (referred to as fully-authentic setup) to the situation where game play takes place in smaller groups lead by researchers as instructors (semi-authentic setup). Instructors own enthusiasm in dealing with material may affect students' absorption of values and importance [15].

2 Related Mathematic Game Studies

Criticisms have been raised towards game studies that either are non-authentic in their setup [5] or lack control groups to the treatment [4]. Therefore, we restrict related work to longer experimental studies in authentic settings where both achievement and affective measures for math are investigated and compared to controls.

Ke and Grabowski [16] used strategy games for arithmetic problem-solving and enrolled 125 voluntary 5th grade students in a 4-week study. Three conditions were investigated: cooperative, competitive and no game play. For math performance, both game playing groups performed significantly better than the control. For attitude, the cooperative game play group was significantly more positive than the other two conditions. This study relates to ours by student age, general topic, and two measures.

In a larger (N=358) similar study, Ke also investigated metacognitive awareness by a self-report questionnaire [6]. Metacognition and self-efficacy are related: while self-efficacy is a predictor of both declarative and procedural knowledge, metacognition is only related to procedural knowledge [14]. In this study, significant effect in attitude was found, but neither math performance nor metacognitive awareness showed effects. This study is similar to our, the main difference is the Teachable Agent game.

In a 18-week study with 193 9th and 10th grade students, the effects of a mission-based game for algebra was investigated [5]. The results indicated significant gain in math achievement. No significant improvement was found in the motivation of the groups.

A geometry puzzle game was used in a 10 session study enrolling 29 6th grade students with 2 play conditions: with and without level progression [17]. The authors found significant performance gains for both treatment groups compared to pre tests results and controls. The affective measure reported gains, but concerned attitude towards the game instead of towards math.

Finally, our game has been evaluated in two previous studies: a short-term pilot study [18] and a study involving a small number of special education students [19].

3 The Teachable Agent Arithmetic Game

The educational content of the game is basic arithmetic with a particular focus on conceptual understanding of base-10 and the arithmetic operations. The approach in our environment is to provide 1) an animated, graphical model simulating arithmetic behavior; 2) a set of two-player games based on the model; and 3) intelligent, teachable agents which can be taught to play the games. In the animated simulation model, square-boxes are explicitly packed/unpacked, to illustrate carrying and borrowing. The computation 48+43 = 91 is illustrated in Fig. 1:

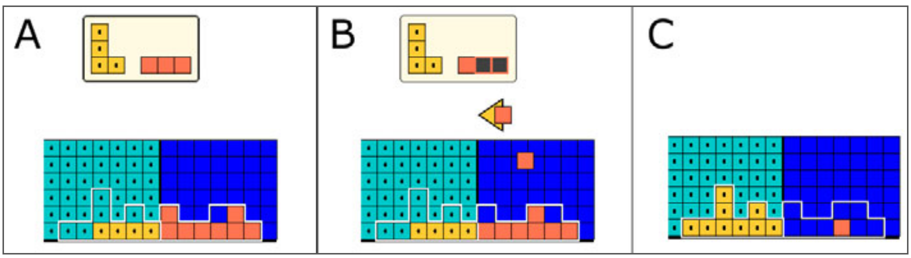

Fig. 1. The carrying operation as an animated packing of squares into a square box

In A, the number 48 is represented graphically on the game board at the bottom (4 orange one-dot squares in the left compartment, and 8 red squares in the right), and the number 43 on the card above. Addition is to put objects on the board, subtraction to remove. Picture B captures the animated packing of 10 red singleton squares into a sized 10 square box. In C, the computation 48+43 = 91 is completed.

In the games, each player acts an arithmetic operation and receives a set of cards with graphical numbers. The players take turn choosing a card until all cards are played. A game, i.e., a sequence of turns, thus constitutes a computation $x_1+y_1+x_2+$.... The player's task is for each turn to choose the best possible card according to various game goals, such as maximizing number of carryings or number of zeroes in the intermediate results. The task involves predicting the cards effects (i.e., the results of one-step computations), reasoning about the available choices for the current turn, and longer term strategies to maximize scoring. The target knowledge is structural properties of the base-10 system, and how numbers behave under computations.

Besides playing themselves, students can teach an agent to play the game in a master-apprentice manner. Students take on the role as teacher, which most find very engaging. Agents are taught in two ways; by showing how to play or by having the agent try making a choice according to its knowledge, which the student either accepts or corrects to a (possibly) better choice. Either way, the agent asks the student reflective questions *on why* a particular choice was made or was better. For example, in figure 2, Mike is teaching his agent in show mode, and has just chosen the card 39, instead of the other choices 33, 97 and 40. The agent, being an inquisitive learner, asks Mike why the choice 39 was good before the computation takes place and the effect is known. The system also provides plausible explanatory responses (with one correct explanation) in a multiple-choice format, for the student to choose from. The agent learns from observing and analyzing the students playing behavior, and from the question responses.

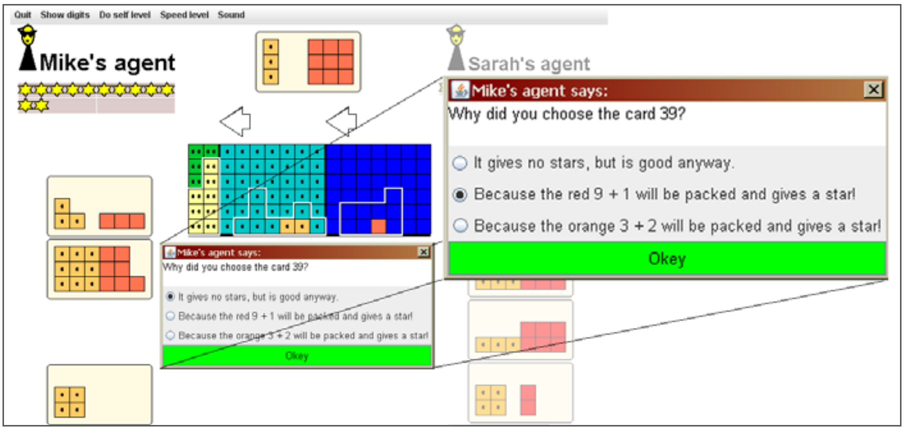

Fig. 2. Mike is teaching his agent by showing how to play and answer explanatory questions

In this way, the TA provides guidance to connect to symbolic math and stimulates reflection of game playing behavior, often required in games to help learners achieve deep understanding [20], [3]. Reflective thinking is an important condition for learning mathematical concepts [17]. The game design adheres to pedagogical fidelity by promoting reflective thinking, allowing exploration and manipulation of virtual objects; to mathematical fidelity by ensuring mathematical soundness in its behavior and by providing substantial training in reasoning and logic; and finally to cognitive fidelity by a its concrete, visual and explicit representations of numbers and operations. The essence of mathematics involves observing and investigating patterns and relationships between objects [7].

4 Method

This study used a pre-post experimental design. The objective of the present study was to evaluate the TA Arithmetic game with respect to the hypotheses that playing the game would: 1) support students' conceptual understanding of basic arithmetic as revealed in the difference in post-test vs. pre-test scores on a mathematical comprehension paper-and-pencil test; 2) scaffold more positive attitudes in students towards the topic of mathematics as revealed in the difference in post-test vs. pre-test scores on an attitudes questionnaire; and 3) scaffold better self-efficacy beliefs as revealed in the difference in post-test vs. pre-test scores on a task-specific self-efficacy questionnaire. We also explored in a secondary analysis if students' self-reported like/dislike for mathematic and/or authenticity levels affected achievements.

The study enrolled five 3rd grade classes and four 5th grade classes at two different locations, southern and west of Sweden, in total 153 students. Due to school policy and practical reasons for a longer in-class, within curriculum study, we had to split conditions at the level of class. The enrolled classes were chosen, at each location, to be as similar as possible with respect to socio-economic background, overall performance, digital competence, amount of math instruction, and pedagogical approach. One class from each location and level were assigned playing condition, the others a no-intervention control condition. Each year, all 3rd and 5th grade students in Sweden take mandatory standard tests in mathematics at a pre-determined period of some weeks, which occurred during the study period. Much attention is paid under math classes to prepare and take the tests during this time, since results are basis for nationwide quality comparisons. This ensured that the math activities during the study were as equivalent as possible between conditions, apart from the intervention.

One of the locations had a semi-authentic setting (game play in groups of 8, monitored by researchers) the other a fully authentic setting (entire class, their regular teachers). There were 68 students in the play condition, 51,5% girls and 48,5% boys.

The playing classes used the game for 9 weeks, aiming for one 40-minute session per week, instead of other activities during regular mathematics classes. Control conditions proceeded with regular instruction. Prior to the study both conditions completed a paper and pencil test in three parts: 1) arithmetic base-ten math problems, 2) questions regarding general attitudes to mathematics, and 3) questions addressing math self-efficacy. After the intervention both conditions completed a post-test with the same three parts and questions as the pre-test. Repeated measures were collected, but only as in-game progression parameters, which are yet to be analyzed.

The math test consisted of 36 items in 7 problem types, adapted to the two age-groups. Several problems were inspired by previous standard national tests, for example using alternative representations of the base-10 system such as "nature-money" where leaves, cones and stones represent ones, tens and hundreds, respectively. Tasks included translating between nature-money and integers, using nature-money for computations and judge the value of nature objects (place value). Other tasks involved deciding which of two sums is the greatest (e.g., *857+275* or *475+639*) or deciding if a sum's result will be an even ten (e.g., 361+439) by reasoning rather than performing formal computations. Students were told they did not need to calculate; some of the examples were deliberately too difficult for them to compute, and there were no room for calculations. To answer such problems by reasoning require a rather deep understanding of the base 10 system and addition.

The questionnaire for assessing attitudes towards math was inspired by Bandura's design guidelines for self-efficacy scales. For general attitude, one explicit ("*Do you think math is boring or fun*?" on a continuous scale very boring to very fun) and 4 implicit questions of math liking such as *"What do you think about learning new topics in math?* were included. The explicit and implicit questions will be correlation tested for validation. For self-efficacy 4 items such as "*How confident are you in deciding which of the sums 47+32 or 35+41 is the largest?*, which all were task-specific in the sense that concrete examples are given for confidence judgments.

Since a Likert scale format with a mark for neutral was used, affective responses were measured with a ruler on the scale (-3,3) with 0 as the neutral point, resulting in the range of (-12,12) for attitude and self-efficacy questions. The explicit like/dislike question (hereafter called math enjoyment variable) was asked only once, since such opinion is considered as a stable property [21] and used for categorization into low, medium and high positive attitude towards mathematics.

5 Results

First, we present the descriptive statistics for the pre test results in Table 1 below:

Table 1. Descriptive statistics of Pre tests for the play and the control group

		Pre Math Achievement (max 36)			Pre Attitude (min -12, max 12)			Pre Self-Efficacy (min -12, max 12)		
	N	Mean	Median	SD	Mean	Median	SD	Mean	Median	SD
Control	**85**	26,11	27,00	5,51	2,67	2,65	4,90	6,79	8,07	4,19
Play	**68**	24,86	26,00	7,02	3,97	4,00	4,25	5,50	6,16	4,39
Total	**153**	25,55	27,00	6,24	3,25	3,50	4,66	6,22	6,98	4,31

A pre-treatment test (Mann-Whitney) was conducted for between group comparisons. The results shows that there are no significant pre-treatment differences between the two condition groups, neither for math achievement, general attitude nor self-efficacy (all *p> ,05*). Neither is there any group difference of the math enjoyment indication, which correlates strongly to the attitude measure. Hence, we can compare the score gain directly, as shown in Table 2 (descriptive statistics).

Table 2. Descriptive statistics of Gain (difference pre and post test)

		Gain Math Achievement			Gain Attitude			Gain Self-Efficacy		
	N	Mean	Median	SD	Mean	Median	SD	Mean	Median	SD
Control	**85**	,93	1,00	4,38	1,50	1,83	3,82	-,23	,00	3,62
Play	**68**	3,19	3,00	5,28	,93	,50	3,09	1,44	1,20	3,52

To examine our three main hypothesis, we have conducted a Mann-Whitney between group comparison test for score gains, showing that there are significant effects for the math achievement (*p=,01*) and the self-efficacy (*p=,009*) in favor of the treatment group, but no significant effect for the attitude gain score *(p=,172).* Effect sizes are *0,47* for achievement and self-efficacy, and *-0,16* for attitude. For comparison, an ANCOVA controlling for pre-test results yields similar p-values for math achievement and self-efficacy (*p=,01* and *p=,03*) and considerably higher for attitude (*p=,67*). Hence, hypothesis 1 and 3 are supported, but not 2.

The explanatory secondary analysis of the within treatment group difference with respect to the categorization of the self-reported math enjoyment variable showed math achievement mean gain for the respective sub groups *(low=9,00; medium=3,14; high=2,35; total=3,19)*, for the general attitude measure *(low=1,23; medium=0,44; high=1,16; total=0,93)*, and for the self-efficacy measure *(low=3,29; medium=0,09; high=1,94; total=1,44)*. There are no significant differences between the subgroups in any of the three measures. For the categorization into semi- and fully authentic groups, there was a slightly larger gain for the semi-authentic group for math achievement (*3,41* compared to *2,97,n.s.*), significantly larger gain in attitude for the semi-authentic group (*1,23* compared to *0,14, p=0,026*), and finally a slightly larger gain of self-efficacy for the fully authentic group (*1,77* compared to *1,14, n.s.*).

6 Discussion and Conclusion

The results support the hypothesis that playing the game improves students' conceptual arithmetic understanding and increases students' self-efficacy beliefs, but not the hypothesis that it scaffolds more positive attitudes towards mathematics in general. For math performance, similar findings using other games are reported in [17], [16], and [5], and the present result strengthens previous indicative results [18].

Measuring attitude change in relation to game usage seems to be a more diverse issue, both in terms of used measures and what the results indicate. Three related studies showed a positive change in attitude, whereas this study and [5] did not. The overall attitude measure used in [6], where a positive change was detected, ought to be compared to both our measures attitude and self-efficacy, since we separated the issues whereas Ke included both in one measure. The attitude measured in [17] is not comparable, since it concerned attitude towards the game and not the subject mathematics. Considering the students' positive engagement and attitude towards the game (as evident from observations, teachers' and students' testimonies), the lack of attitude gain in our study may be explained by students not including the game play in general mathematics did (as indicated in post intervention interviews) and that 9 weeks is too short to change an attitude formed during several years [21].

We consider the positive change in self-efficacy beliefs to be the main contribution of this paper since it has not been studied as a separate issue in related works. Also, self-efficacy beliefs are strong predictors of future math accomplishments [12]. We suggest that the particular game design contributes to an explanation by: 1) the absence of failures; choices can be better or worse but never wrong, and 2) the TA allowing students to act the role of an expert, boosting self-esteem and confidence. A future study comparing the game with and without the TA, should shed light on 2).

The exploratory analysis of within treatment group difference with respect to math enjoyment and authenticity level were not significant but give indications for further research. (For example, a future larger study allowing multi-level analysis, e.g., hierarchical linear model analysis, should provide further insights.) The low attitude group show the largest gains on all three measures, which may suggest that an unconventional approach to math could particularly attract these students (also observed in [19]). Being at-risk and a difficult group to attract, such indications deserve further research. Finally, we can only speculate on why the significant difference in attitude gains between the semi- and fully-authentic groups (in favor of semi-authentic) appeared, but perhaps the extra attention, the authority of being researchers or their conviction of the game's value played a role.

References

1. Rieber, L.P.: Seriously considering play: Designing interactive learning environments based on the blending of microworlds, simulations, and games. Educational Technology Research & Development 44(2), 43–58 (1996)
2. Schwartz, D.L., Martin, T.: Inventing to Prepare for Learning: the Hidden Efficiency of Original Student Production in Statistics Instruction. Cognition and Instruction 22, 129–184 (2004)
3. Moreno, R., Mayer, R.E.: Role of guidance, reflection and interactivity in an agent-based multimedia game. Journal of Educational Psychology 97(1), 117–128 (2005)
4. Vogel, J.F., Vogel, D.S., Cannon-Bowers, J., Bowers, C.A., Muse, K., Wright, M.: Computer gaming and interactive simulations for learning: A meta-analysis. Journal of Educational Computing Research 34(3), 229–243 (2006)
5. Kebritchi, M., Hirumi, A., Bai, H.: The effects of modern mathematics computer games on mathematics achievement and class motivation. Comp. & Education 55, 427–443 (2010)
6. Ke, F.: Computer games application within alternative classroom goal structure: cognitive, metacognitive and affective evaluation. Educational Teachnology Research Development 56, 539–556 (2008)
7. Bos, B.: Virtual math objects with pedagogical, mathematical, and cognitive fidelity. Computers in Human Behavior 25, 521–528 (2009)
8. Biswas, G., Katzlberger, T., Brandford, J., Schwartz, D.L., TAG-V.: Extending intelligent learning environments with teachable agents to enhance learning. In: Moore, J.D., Redfield, C.L., Johnson, W.L. (eds.) Artificial Intelligence in Education, pp. 389–397 (2001)
9. Schwartz, D.L., Chase, C., Wagster, J., Okita, S., Roscoe, R., Chin, D., Biswas, G.: Interactive Metacognition: Monitoring and Regulating a Teachable Agent. In: Hacker, D.J., Dunlosky, J., Graesser, A.C. (eds.) Handbook of Metacognition in Education (2009)
10. McLeod, D.B.: Research on Affect and Mathematics Learning in the JRME: 1970 to the Present. Journal for Research in Mathematics Education 25(6), 637–647 (1994)
11. Ma, X., Kishor, N.: Attitude Toward Self, Social Factors, and Achievement in Mathematics: A Meta-Analytic Review. Educational Psychology Review 9(2), 2 (1997)
12. Pajares, F., Graham, L.: Self-Efficacy, Motivation Constructs, and Mathematics Performance of Entering Middle School Students. Contemporary Educational Psychology 24, 124–139 (1999)
13. Bandura, A.: Self-efficacy: The Foundation of Agency. L. Erlbaum, Mahwah (2000)
14. Moores, T.T., Chang, J.C.-J., Smith, D.K.: Clarifying the role of self-efficacy and metacognition as predictors of performance: construct development and test. SIGMIS Database 37(2-3), 125–132 (2006)
15. Frenzel, A.C., Pekrun, R., Goetz, T.: Perceived learning environment and students' emotional experiences: A multilevel analysis of mathematics classrooms. Learning and Instruction 17, 478–493 (2007)
16. Ke, F., Grabowski, B.: Gameplaying for maths learning: cooperative or not? British Journal of Educational Technology 38(3), 249–259 (2007)
17. Sedig, K.: Toward operationalization of 'flow' in mathematics Learnware. Computers in Human Behavior 23, 2064–2092 (2007)
18. Pareto, L., Schwartz, D.L., Svensson, L.: Learning by guiding a teachable agent to play an educational game. In: Proc. of the 14th Int. Conference on Artificial Intelligence in Education, pp. 662–664. IOS press, Amsterdam (2009)

19. Nilsson, A., Pareto, L.: The Complexity of Integrating Technology Enhanced Learning in Special Math Education – A Case Study. In: Wolpers, M., Kirschner, P.A., Scheffel, M., Lindstaedt, S., Dimitrova, V. (eds.) EC-TEL 2010. LNCS, vol. 6383, pp. 638–643. Springer, Heidelberg (2010)
20. Mayer, R.E.: Should there Be a Three-Strikes Rule Against Pure Discovery Learning? The Case for Guided Methods of Instruction. Educational Psychologist 59, 14–19 (2004)
21. Middleton, J.A., Spanias, P.A.: Motivation for Achievement in Mathematics: Findings, Generalizations, and Criticisms of the Research. Journal for Research in Mathematics Education 30(1), 65–88 (1999)