Learning and Individual Differences 24 (2013) 190–197

Contents lists available at SciVerse ScienceDirect

Learning and Individual Differences

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

The influence of experiencing success in math on math anxiety, perceived math competence, and math performance

Brenda R.J. Jansen [a,*], Jolien Louwerse [a], Marthe Straatemeier [b], Sanne H.G. Van der Ven [b], Sharon Klinkenberg [b], Han L.J. Van der Maas [b]

[a] Department of Psychology, Developmental Psychology, University of Amsterdam, The Netherlands
[b] Department of Psychology, Psychological Methods, University of Amsterdam, The Netherlands

ARTICLE INFO

Article history:
Received 2 May 2012
Received in revised form 2 November 2012
Accepted 22 December 2012

Keywords:
Computer-adaptive practice
Math anxiety
Perceived competence
Math performance
Arithmetic

ABSTRACT

It was investigated whether children would experience less math anxiety and feel more competent when they, independent of ability level, experienced high success rates in math. Comparable success rates were achieved by adapting problem difficulty to individuals' ability levels with a computer-adaptive program. A total of 207 children (grades 3–6) were distributed over a control and three experimental conditions in which they used the program for six weeks. Experimental conditions differed in pre-set success rate. Math anxiety, perceived math competence, and math performance were assessed before and after the practice period. Math anxiety scores improved equally in all conditions. Improvement on perceived math competence was modest. Math performance, however, only improved in the experimental conditions. Moreover, the higher the pre-set success rate, the more problems were attempted, and the larger the improvement in math performance, suggesting that success in math leads to more practice and thus to higher math performance.

1. Introduction

Despite digital resources, math skills remain undeniably important in everyday life (e.g., Ashcraft & Krause, 2007). In elementary school, there is ample attention for teaching those skills. However, the development of math skills may be threatened by math anxiety (e.g., Hembree, 1990) and feelings of incompetence (Harter, 1982).

Math anxiety has been associated with negative feelings surrounding math and adverse outcomes on both math performance and the confidence to learn math (Bekdemir, 2010; Chinn, 2009; Chiu & Henry, 1990; Hembree, 1990; Ho et al., 2000; Ma, 1999; Rubinsten & Tannock, 2010). It may be the case that math anxiety causes low math performance because math anxiety and worrisome thoughts tax working memory to such an extent that working memory cannot be optimally used when performing math-related tasks (Ashcraft & Krause, 2007; Krinzinger, Kaufmann, & Willmes, 2009). Math anxiety may also cause low performance because people with high math anxiety tend to avoid math-related tasks (Chinn, 2009; Hembree, 1990) and finish these tasks quickly but inaccurately to end the stressful situation as soon as possible (Ashcraft & Faust, 1994). However, Ma and Xu (2004) found evidence that supports the opposite direction, that low math performance causes later math anxiety. As results of most studies support both directions (Hembree, 1990; Newstead, 1998), a reciprocal model of math anxiety and math performance seems most plausible (see also Van der Maas et al., 2006).

Perceived competence is closely related to test anxiety (Bandalos, Yates, & Thorndike-Christ, 1995; Lee, 2009). Studies show a reciprocal relation between perceived competence and academic performance (Guay, Marsh, & Boivin, 2003; Marsh & Martin, 2011; Valentine, DuBois, & Cooper, 2004). However, children perceive their competence in different domains, such as sports and academic skills, differently (Harter, 1982; Marsh & Martin, 2011; Valentine et al., 2004). Hence, a specific scale that is focused on the math domain is desired.

In the present study, we aim to disentangle the influence of success in math, independent of actual math ability, on math anxiety and perceived math competence. Relatively high success rates in math are ensured for all children by adapting the difficulty level of problems to children's individual ability levels. Practicing at the individual ability level is possible in computer-adaptive programs. The specific computer-adaptive program that was used is Math Garden (Klinkenberg, Straatemeier, & van der Maas, 2011), which applies a new type of Computer Adaptive Testing (CAT). In both original CAT and Math Garden, the individual ability level is estimated dynamically, that is, while an individual solves problems. Selection of problems depends on the estimation of an individual's ability level and on a preset level of expected success rate.

Math Garden differs from original CAT in the height of the preset level of expected success rate. In original CAT, an expected success

* Corresponding author at: Weesperplein 4, 1018 XA Amsterdam, The Netherlands. Tel.: +31 20 5256735; fax: +31 20 6390279.
E-mail addresses: b.r.j.jansen@uva.nl (B.R.J. Jansen), M.Straatemeier@uva.nl (M. Straatemeier), S.H.G.vanderVen@uva.nl (S.H.G. Van der Ven), S.Klinkenberg@uva.nl (S. Klinkenberg), H.L.J.vanderMaas@uva.nl (H.L.J. Van der Maas).

1041-6080/$ – see front matter © 2013 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.lindif.2012.12.014

rate of 50% is used for the selection of problems, as problems in this range are maximally informative in estimating individuals' abilities. Math Garden allows for higher expected success rates (60% and above) because response time is used as an additional source of information on individuals' ability levels. We hypothesized that when children experience that they can solve the majority of math problems, their math anxiety may lower and perceived math competence may increase. Given the reciprocal relation between emotional experience of math and math performance, improvement of math performance is expected as well. Moreover, three levels of success rate are used (60%, 75%, and 90%) and effects of this manipulation on math anxiety, perceived math competence, and math performance are investigated. Other aspects of Math Garden might be beneficial as well. These features: game character of the program, reward of speed, colorful environment, and immediate feedback, however, were not manipulated.

An additional advantage of solving math problems on a computer is the reduction of the number of public embarrassing math-related experiences in the classroom. Negative experiences with math appear to contribute to math anxiety (Rubinsten & Tannock, 2010), especially when these are public (Bekdemir, 2010; Newstead, 1998).

Summarized, we investigated whether practicing math with a computer-adaptive program had beneficial effects on math anxiety, perceived math competence, and math performance. Beneficial effects were expected because the program offered high success rates, because problems were adapted to the individual level, and because errors were made in private. In addition, we hypothesized that increasing success rate was associated with decreasing math anxiety, increasing perceived math competence, and increasing math performance.

Primary school children were randomly assigned to four conditions. In the control condition, children practiced math as usual. In the three experimental conditions, children practiced math with the computer-adaptive program for six weeks. Mean targeted success rates in the experimental conditions were set at 60%, 75%, and 90%. Math anxiety, perceived competence, and math performance were assessed before and after the practice period. In all analyses, we investigated the relation with grade because it is found that math anxiety increases with age (Hembree, 1990) and because absolute math performance increases with age (Krinzinger et al., 2009; Van der Ven, Kroesbergen, Boom, & Leseman, 2012). We also investigated the relation with gender. Results concerning gender differences in math anxiety vary across studies. Gender differences (with females having a higher level of math anxiety) were found in a Taiwanese sample of sixth-graders (Ho et al., 2000), but only small gender differences were found in a UK sample of 11–17-year-olds (Chinn, 2009) and no gender differences were found in Chinese and American sixth-graders (Ho et al., 2000), an American sample of sixth- and seventh-graders (Chiu & Henry, 1990), an American sample of seventh-graders (Ma & Xu, 2004) and a meta-analysis (Ma, 1999). Possibly, gender differences in math anxiety differ between cultures and between age groups. Gender differences in math performance are debated. A large-scale study of 15-year-olds shows that the male advantage in math is small but consistent across domains (Liu & Wilson, 2009) although the gender gap seems to narrow in recent decades (e.g., Bonnot & Croizet, 2007).

2. Method

2.1. Participants

A total of 252 children participated, ranging in age from 8 to 13 years. A number of children were not present at the posttest because it took place shortly before the summer holidays and the children left earlier because of holidays. Other children were absent due to illness. Altogether, 39 children were absent from one of the measurements. Data from six children in experimental conditions were excluded because they had attempted fewer than seven problems in the computer-adaptive program. The final sample consisted of 207 children.

Participants came from two primary schools in The Netherlands. In the largest school ($N = 192$), 62% of the students were at risk of falling behind in their education[1]; in the smallest school ($N = 60$), the percentage of students at risk was 31% (Source: Ministry of Education, Culture, and Science, The Netherlands). Parents of both schools were informed and could refuse participation of their child. However, none of the parents refused. The Ethics Committee of the Psychology Department approved of the procedures.

We placed children randomly into one of four conditions, stratified on the basis of age, gender, and pretest scores: a control condition and three experimental conditions, “Difficult” (60% correct), “Medium” (75% correct), and “Easy” (90% correct). Children from various conditions were in one classroom. Table 1 shows number of participants, average age, gender distribution, and grade distribution by condition. An ANOVA indicated that conditions did not differ in age, $F(3, 206) = 0.20$, $p = .893$, or gender, $\chi^2(3) = 1.96$, $p = .582$.

2.2. Design and procedure

The study started in the spring of 2011. The two schools started two weeks apart. Before the start, teachers received an informative letter, which stated that participating children would practice in Math Garden for 3 to 5 times per week and that a session would take about 10 to 15 min.

Participants completed a pretest and a posttest. In the pretest, perceived competence, math anxiety, and math performance were sequentially assessed. All tests were paper-and-pencil and administration was group-wise. Administration lasted about 1 h. Children in all conditions followed the regular math curriculum. Children in the experimental conditions used Math Garden as well.

The pretest took place just before a holiday leave. After this leave, a trained graduate student explained Math Garden to the children in the experimental conditions and their teachers. She distributed log-ins and passwords to the children and provided the teachers with lists to track their pupils' playing frequency. The practice period lasted six weeks. In this period, the student regularly visited the schools. Integration of Math Garden in the curriculum differed between teachers, and regulations regarding the curriculum were absent. Math Garden tracked playing frequency automatically and teachers received an update of their pupils' playing frequency twice. Weekly practice was recommended but was not registered with sufficient frequency to check whether this was accomplished.

Because of holidays, the posttest was not conducted immediately after the practice period but a few weeks later. It included the exact same measurements of perceived competence, math anxiety and math performance. On average, there were 11.1 weeks ($SD = 0.36$ week) between pre- and post assessments. Finally, participants were debriefed and told that success rates had varied. Participants in the control condition received their log-ins and passwords in this final session.

2.3. Materials

2.3.1. Math anxiety

The Math Anxiety Scale for Children (MASC; Chiu & Henry, 1990) was translated into Dutch. The original MASC consists of 22 items, each describing a situation that concerns math (e.g., “Being given a math quiz that you were not told about”). Participants indicated their degree of anxiety in the given situation on a four-point-scale. In the translation (the MASC-NL), one situation was removed because it was uncommon for the Dutch school system (“Using the tables in the back

[1] Risk of falling behind is determined by combining the educational level of the child's parents and the average income of the residents in the child's home district (Source: Ministry of Education, Culture, and Science, The Netherlands).

Table 1
Characteristics of participants by condition.

	Condition				
	Difficult	Medium	Easy	Control	Total
N	51	52	48	56	207
Mean age in years	10.8 (1.3)	10.7 (1.3)	10.9 (1.3)	10.8 (1.3)	10.8 (1.3)
% males	61%	54%	48%	50%	53%
N in grade 3	15	15	10	12	52
N in grade 4	11	11	11	14	47
N in grade 5	12	15	16	16	59
N in grade 6	13	11	11	14	49

Note. Standard deviations are between parentheses.

of a math book”). To compensate, two situations were added (“You need to discover the math problem in a story”; “You don't understand a math problem”). In total, the MASC-NL consisted of 23 items. Scores, therefore, ranged from 23 to 92, with a higher score indicating a higher level of math anxiety. Cronbach's alpha was .93 for the MASC-NL pretest scores. Pearson's correlation coefficient between the MASC-NL and math performance (for instrument: see Section 2.3.3) was $r = -.22$, $p = .002$ on the pretest. Pearson's correlation coefficient between pre assessment and post assessment of the MASC-NL was $r = .89$, $p < .001$, for participants in the control condition. This correlation may be conceived as an indication of test–retest reliability.

2.3.2. Perceived math competence

Scales “Cognitive Competence”, “Social Competence”, and “General Self-worth” from the Perceived Competence Scale for Children (Harter, 1982) were administered. The Dutch translation (Veerman, Straathof, Treffers, Van den Bergh, & Ten Brink, 1997) was used and extended with the scale “Math Competence”. Each scale consisted of six pairs of statements. An example of a pair of statements was “Some children are good at math” and “Other children have a bit more trouble with math”. Participants selected the statement that applied most to them. Next, participants indicated whether the statement was “completely true” for them or was “somewhat true” for them. Hence, responses were scored on a four-point scale, with higher scores indicating higher perceived competence. Cronbach's alpha was .86 for Math Competence pretest scores. Pearson's correlation coefficient between perceived math competence and math performance (for instrument: see Section 2.3.3) was $r = .44$, $p < .001$ on the pretest.

2.3.3. Math performance

TempoTest Automatiseren (TTA; De Vos, 2010) was used to assess the degree of memorization of math facts. Contents of the TTA are somewhat similar to those of Math Garden as both consist of mere sums. However, the TTA is an independent instrument, often used in primary school to monitor children's math performance (grades 3–6). One might argue that insight into math problems is more important than memorization (like learning basic math facts). However, memorization is essential to solve insight problems. Norms for TTA are available for grades 3–6 and differentiation between children should still be possible in the oldest age group. The test consisted of four parts: addition, subtraction, multiplication, and division. Each part consisted of 50 problems and participants were allowed to work on each part for exactly 2 min. The number of problems solved correctly was obtained for each part and for the total number of problems.

2.3.4. Computer-adaptive program

Math Garden (Klinkenberg et al., 2011) is a web based computer-adaptive application for practicing mathematical skills. Central in Math Garden is the child's personal garden (Fig. 1A), containing six flower-beds that correspond to games for addition, subtraction, multiplication, division, fractions, and a Number game (i.e., a version of the “24 game”). Availability of games depended on each individual's ability level. In the Number game, numbers and operation signs need to be combined to obtain a certain target number (e.g., numbers 4, 5, and 1 and signs “×” and “−” needed to be combined to obtain 15 as follows: $(4-1)\times5$).

A single game consisted of fifteen problems presented sequentially (see Fig. 1B for an example of a problem). The screen showed a problem, a response area (six response options in the addition and subtraction games and a “fill-in-the-blank” with a number pad in the other games), a question mark that could be clicked in case a participant did not know the answer, a money bag, a row of coins, and a green bar indicating the number of items left to answer. The time limit for problems was set to 20 s. After a correct answer children were rewarded a virtual coin for every second that was still left, but they lost this number of coins from their collection if the answer was incorrect. Rewards for correct responses and punishment for quick guesses should encourage fast responses when one is certain and discourage guessing when one is uncertain. Maris and Van der Maas (2012) provide the mathematical basis and statistical response model for this scoring rule. The correct answer was shown after every problem. Coins could be exchanged for virtual prizes.

After playing a game, children automatically returned to their garden. If the child did well, plants in the corresponding flower-bed had grown. Children could play both during school time and at home because Math Garden is web-based.

Selection of problems was regulated by an adaptive algorithm (Klinkenberg et al., 2011) and depended on the estimated skills of the participant, estimated problem difficulties and experimental condition. Problems that a participant could solve correctly with a pre-set average percentage correct of 60%, 75%, and 90% were selected for conditions Difficult, Medium, and Easy, respectively. Since the ability estimate of the participant was updated after every problem, these percentages could be maintained throughout the study, regardless of the level and/or the progress of the child. These preset percentages are only reached if the adaptive algorithm operates optimally. The observed percentages correct are reported in the Results section.

3. Results

To indicate effect size, Cohen's *d* was reported in case of *t*-tests. In the case of an ANOVA, concerning a comparison of means of more than two groups, *r* was used because Cohen's *d* can only be applied for comparing means of two groups (Rosenthal, Rosnow, & Rubin, 2000).

3.1. Manipulation checks

Table 2 shows average success rate and number of problems attempted by condition. Success rates were 66%, 72%, and 81% for conditions Difficult, Medium, and Easy, respectively. Levene's test was performed to test homogeneity of variances of success rates across conditions and was not significant, $F(2, 148) = 2.84$, $p = .062$. One-sample *t*-tests, comparing the average success rates to the aimed success rates, showed that success rate in condition Difficult was significantly higher than 60%, $t(50) = 5.05$, $p < .001$, $d = .70$. Success rates in conditions Easy and Medium were significantly lower than the aimed rates, Easy: $t(47) = -9.44$, $p < .001$, $d = 1.37$; Medium: $t(51) = -3.37$, $p = .001$, $d = .47$. Nevertheless, an ANOVA on success rates, with experimental condition as factor, revealed that, as intended, experimental conditions differed significantly in average success rate, $F(2, 150) = 60.06$, $p < .001$, $r = .67$. In this analysis, experimental condition was treated as an ordinal variable, increasing with increasing success rate (Difficult = 1, Medium = 2, and Easy = 3). A linear trend was significant, $F(1, 150) = 119.31$, $p < .001$, $r = .67$, indicating that increase of success rate was associated with a higher percentage correct.

The number of attempted problems reflects the total frequency of practice in the classroom and outside school hours. Levene's test was performed to test homogeneity of variances of number of problems attempted across conditions and was not significant, $F(2, 148) = 1.08$,

Fig. 1. Screen shots of Math Garden. 1A: Personal garden. 1B: Subtraction problem.

Table 2
Average success rate, number of problems attempted, and scores on measurements of math anxiety, perceived math competence, and math performance, by condition.

	Condition			
	Difficult	Medium	Easy	Control
Success rate				
	66% (8%)	72% (6%)	81% (7%)	–
Number of problems attempted				
	423.8 (334.2)	548.1 (375.2)	645.7 (390.2)	–
Math anxiety				
Pre	39.0 (10.5)	42.5 (14.0)	38.2 (12.0)	40.0 (13.4)
Post	37.1 (10.6)	37.1 (10.1)	34.8 (10.3)	38.0 (13.2)
Perceived math competence				
Pre	17.8 (4.3)	16.9 (4.9)	17.0 (4.6)	17.5 (4.8)
Post	18.1 (4.2)	17.8 (4.0)	17.6 (4.5)	17.5 (4.3)
Math performance: total				
Pre	136.0 (32.8)	131.4 (37.9)	137.7 (36.6)	143.6 (36.5)
Post	142.2 (33.7)	140.4 (39.5)	149.3 (35.6)	144.6 (37.6)
Math performance: addition–subtraction scores				
Pre	68.4 (14.1)	68.0 (15.5)	67.7 (16.4)	71.9 (13.4)
Post	71.3 (15.8)	71.8 (16.3)	73.9 (15.9)	73.5 (15.4)

Note. Standard deviations are between parentheses.

$p = .343$. A regression analysis with number of problems attempted as dependent variable and condition as independent variable (experimental conditions only; Difficult = 1, Medium = 2, and Easy = 3) was performed to investigate whether condition was associated with number of problems attempted. Condition was a significant predictor, $F\,(1, 150) = 9.13$, $p = .003$, $r = .24$; $B = 111.08$, indicating that the higher the success rate, the more problems were attempted. The number of problems attempted did not differ between boys and girls, $t\,(150) = -1.11$, $p = .27$, $d = .18$.

Table 2 also shows average scores on measurements of math anxiety, perceived math competence, and math performance, for both pretest and posttest, by condition. First, we compared scores of the four conditions on pretest measurements. Levene's test was significant for pretest math anxiety scores, $F\,(3, 203) = 2.90$, $p = .046$, indicating that variances were heterogeneous. Therefore, scores were compared by means of the Kruskal–Wallis test. The Kruskal–Wallis test was not significant, $H\,(3) = 2.92$, $p = .405$, indicating that conditions did not differ in pretest math anxiety scores. Levene's test was not significant for pretest scores of perceived math competence and math performance, $F\,(3, 203) = 0.67$, $p = .574$ for perceived math competence; $F\,(3, 203) = 0.62$, $p = .601$ for math performance. One-way

ANOVAs with condition as factor and measurements of perceived math competence and math performance as dependent variables, respectively, showed that conditions also did not differ significantly on these pretest measurements (all p's ≥ .370). In subsequent analyses, we focused on difference scores: the difference between scores on the pretest and the posttest. Fig. 2 displays average difference scores by condition for math anxiety, perceived math competence, and math performance. A negative difference corresponded to reduced math anxiety whereas a positive difference corresponded to improvement for perceived math competence. With respect to math performance, we focused on the sums of addition and subtraction scores (theoretical range: 0–100) because multiplication and division scores appeared to suffer from a ceiling effect. The ceiling effect for multiplication and division scores seemed to occur in grades 4–6, at both pretest and posttest. Correlations between addition, subtraction, multiplication, and division scores were, however, high: on both pretest and posttest, r was between .62 and .87, all p's < .001. A positive difference indicated improvement.

3.2. Math anxiety

First, grade and gender differences on pretest and posttest scores were considered to decide whether these factors should be included in the analysis of difference scores. For gender, Levene's test was not significant for both pretest and posttest. However, for grade, Levene's test was almost significant for pretest math anxiety scores, F (3, 203) = 2.55, p = .057, and significant for posttest math anxiety scores, F (3, 203) = 4.00, p = .009, indicating that variances of grades were heterogenous. Hence, t-tests for independent samples were used to compare boys and girls, whereas Kruskal–Wallis tests were used to compare different grades.

T-tests for independent samples, with gender as independent variable and pretest and posttest math anxiety scores as dependent variables, respectively, were not significant, showing that girls and boys did not differ on pretest and posttest math anxiety scores, t (205) = −.96, p = .337 for pretest scores; t (205) = −1.58, p = .116 for posttest scores. Kruskal–Wallis tests with grade as independent variable and pretest and posttest math anxiety scores as dependent variables, respectively, were also not significant, indicating that grades did not differ in pretest and posttest math anxiety scores either, H (3) = 2.83, p = .419 for pretest scores; H (3) = 0.80, p = .850 for posttest scores. Grade and gender were therefore not included in subsequent analyses on math anxiety difference scores.

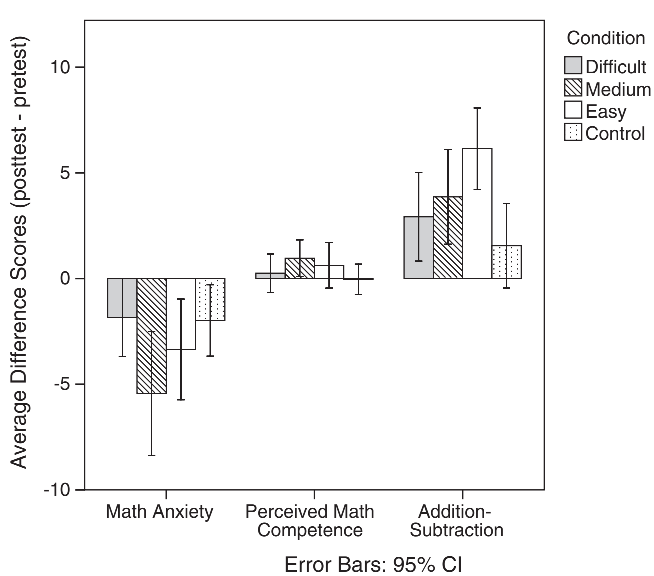

Fig. 2. Difference scores on measurements of math anxiety, perceived math competence, and addition and subtraction, by condition. Error bars show 95% confidence intervals.

Fig. 2 shows that the average difference score of each condition was at or below zero. Zero was not in the confidence interval for conditions Medium, Easy, and Control. The upper boundary was at zero for condition Difficult. Hence, reported math anxiety had reduced significantly in all conditions. Levene's test, performed to test homogeneity of variances of math anxiety difference scores, of the four conditions, was significant, F (3, 203) = 4.90, p = .003. Hence, a Kruskal–Wallis test was performed to investigate differences in math anxiety difference scores between conditions. The Kruskal–Wallis test on math anxiety difference scores, with condition as independent variable (four levels: Difficult, Medium, Easy, and control), demonstrated that the main effect of condition was not significant, H (3) = 3.70, p = .296. Summarized, participants in all conditions showed lower levels of math anxiety on the posttest compared to the pretest, but this effect was equal across conditions.

3.3. Perceived math competence

Levene's test was performed to test for homogeneity of variances of perceived competence scores of boys and girls, and of different grades. For gender, Levene's test was significant, both on the pretest, F (1, 205) = 9.03, p = .003, and on the posttest, F (1, 205) = 4.97, p = .027. Also for grade, Levene's test was significant on the pretest, F (3, 203) = 4.06, p = .008, and on the posttest, F (3, 203) = 4.99, p = .002. Hence, Mann–Whitney tests were performed to investigate gender differences and Kruskal–Wallis tests to investigate grade differences on perceived competence scores. Both on pretest and posttest, the Mann–Whitney test was non-significant, U = 4730, p = .168 for pretest; U = 4843, p = .251 for posttest. In addition, the Kruskal–Wallis test was non-significant for both pretest and posttest, H (3) = 2.09, p = .553; H (3) = 2.428, p = .489. These results indicate that there were neither gender differences nor grade differences in pretest and posttest perceived competence scores. Hence, grade and gender were not included in subsequent analyses on perceived competence difference scores.

Fig. 2 shows that the average difference score of each condition was only slightly higher than zero and that zero was in the confidence interval for all conditions except condition Medium. Hence, the difference score for perceived math competence was only significantly higher than zero for condition Medium. Levene's test was not significant, demonstrating that variances of perceived competence difference scores were homogeneous for the four conditions. Differences between conditions were investigated by a one-way ANOVA on the perceived competence difference scores, with condition as independent variable. The ANOVA indicated that the main effect of condition was not significant, F (3, 206) = 0.99, p = .398, r = .12. Summarized, only participants in condition Medium improved from pretest to posttest on perceived math competence and the improvement was very modest.

3.4. Math performance

Levene's test was not significant for gender and for grade, for both pretest and posttest scores. Univariate ANOVAs with grade as factor and addition–subtraction-scores on pretest and posttest as independent variables indicated that addition–subtraction-scores differed significantly between grades, F (3, 206) = 18.42, p < .001, r = .52 for pretest scores; F (3, 206) = 17.69, p < .001, r = .46 for posttest scores. On the pretest, addition–subtraction scores were 58.35, 67.89, 73.83, and 75.96 for grades 3–6, respectively. Posthoc tests indicated that on the pretest, children in grade 3 scored lower than children in subsequent grades, and that children in grade 4 scored lower than children in grade 6. On the posttest, addition–subtraction scores were 62.33, 69.30, 78.20, and 80.06 for grades 3–6, respectively. Posthoc tests indicated that on the posttest, children in grade 3 scored lower than children in grades 5 and 6, and that children in grade 4 scored lower than children in grades 5 and 6. Hence, the older is the child, the higher is the score, which, however, leveled off in the highest two grades. An independent samples *t*-test indicated that girls performed worse than boys on the pretest, $t\ (205) = -2.18$, $p = .030$, $d = .30$, but not on the posttest. The average score on the pretest was 71.19 for boys and 66.72 for girls. Therefore, we included both age and gender in subsequent analyses on addition–subtraction problems.

Fig. 2 shows that the average difference score of each experimental condition was higher than zero and that zero was not in the confidence interval. The average difference score for the control condition was only slightly higher than zero and zero was in the confidence interval. Hence, difference scores were significantly higher than zero for all experimental conditions, but not for the control condition. An ANOVA was performed to investigate differences between conditions on the difference addition–subtraction scores. Independent variables in the ANOVA were condition, gender, and grade. Levene's tests were not significant, indicating that variances of addition–subtraction difference scores were homogeneous across the four conditions, across boys and girls, and across the four grades. The ANOVA showed that the main effect of condition was significant, $F\ (3, 207) = 3.65$, $p = .018$, $r = .20$. A simple contrast, with the control condition as the reference category, showed that participants in condition Easy improved more than participants in the control condition, $p = .001$. Participants in experimental conditions Medium and Difficult did not improve significantly more than participants in the control condition. Posthoc analyses also showed that improvement in condition Easy was significantly higher than improvement in the control condition and that differences between other conditions were not significant. The main effect of gender was significant as well, $F\ (1, 207) = 4.07$, $p = .045$, $r = .12$, indicating that girls improved more than boys did. Neither the main effect of grade nor any interactions were significant.

3.5. Mediation effect of number of problems attempted

Success rate of condition was positively related to the number of problems attempted (see Section 3.1). Possibly, the number of problems attempted mediated the relationship between success rate and degree of progress in math anxiety, perceived math competence, and math performance. This hypothesis was tested by performing mediation analysis for each domain, using scripts for SPSS by Preacher and Hayes (2008). That is, a hierarchical regression analysis was performed for each domain. In step 1, a regression analysis was performed to estimate the total effect of success rate for each domain. Success rate was operationalized by experimental condition, which was again coded as an ordinal variable, increasing with increasing success rate. The dependent variable was the difference between scores on the pretest and posttest on measurements of math anxiety, perceived math competence, or math performance (addition–subtraction-scores), respectively. Gender (male = 0; female = 1) and grade were treated as covariates. The results of step 1 were compared to the results of step 2, in which a second regression analysis was performed to estimate the direct effect of success rate on each domain, after controlling for the effect of number of problems attempted. Table 3 summarizes the results of the hierarchical regression analysis, by domain.

The explained variance of the regression model in step 1, including the total effect of success rate, was not significant for math anxiety, $R^2 = .03$, $F\ (3, 147) = 1.63$, $p = .185$. The inclusion of number of attempted problems as a predictor in step 2 did not explain a significant additional proportion of variance, $\Delta R^2 = .002$, $\Delta F\ (1, 146) = 0.25$, $p = .620$. Table 3 shows that the estimated beta parameter for success rate was not significant in both models.

The explained variance of the regression model in step 1, including the total effect of success rate, was not significant for perceived math competence, $R^2 = .01$, $F\ (3, 147) = 0.59$, $p = .621$. The inclusion of number of attempted problems as a predictor in step 2 did not explain a significant additional proportion of variance, $\Delta R^2 = .01$, $\Delta F\ (1, 146) = 1.37$, $p = .244$. Table 3 shows that the estimated beta parameter for success rate was not significant in both models.

Table 3
Results of regression analyses that estimate total and direct effects of success rate on difference scores.

	B	*SE B*	β
Math anxiety			
Step 1: Model including total effect of success rate			
Success rate	−0.934	0.873	−.087
Grade	1.203	0.631	.155†
Gender	1.095	1.419	.063†
Step 2: Model including direct effect of success rate			
Success rate	−0.827	0.901	−.077
Grade	1.169	0.637	.151†
Gender	1.147	1.427	.066
Number of problems attempted	−0.001	0.002	−.042
Perceived math competence			
Step 1: Model including total effect of success rate			
Success rate	0.168	0.340	.041
Grade	−0.169	0.246	−.057
Gender	0.526	0.553	.079
Step 2: Model including direct effect of success rate			
Success rate	0.070	0.350	.017
Grade	−0.137	0.247	−.046
Gender	0.478	0.553	.072
Number of problems attempted	0.001	0.001	.100
Math performance (addition–subtraction)			
Step 1: Model including total effect of success rate			
Success rate (path c in Fig. 3A)	1.416	0.741	.154†
Grade	0.134	0.536	.020
Gender	2.798	1.204	.187*
Step 2: Model including direct effect of success rate			
Success rate (path c′ in Fig. 3B)	0.968	0.749	.105
Grade	0.281	0.529	.042
Gender	2.581	1.186	.172*
Number of problems attempted (path b in Fig. 3B)	0.004	0.002	.204*

Note.
† $p<.1$.
* $p<.05$.

The explained variance of the regression model in step 1, including the total effect of success rate, was significant for math performance, $R^2 = .07$, $F\ (3, 147) = 3.39$, $p = .020$. The inclusion of number of attempted problems as a predictor in step 2 explained a significant additional proportion of variance, $\Delta R^2 = .04$, $\Delta F\ (1, 146) = 6.29$, $p = .013$. Fig. 3 shows a schematic representation of both models, with the model comprising the total effect in Fig. 3A (step 1), and the model comprising the direct effect, controlling for number of problems attempted in Fig. 3B (step 2). Table 3 shows that the estimated beta parameter for success rate approached significance in the model including the total effect of success rate (path c in Fig. 3A) but was not significant when number of attempted problems was added as a predictor (path c′ in Fig. 3B). Instead, the significant parameter associated with path a (see Fig. 3B; estimated with a regression analysis with the number of problems attempted as dependent variable and success rate as independent variable; see Section 3.1) signaled a positive relation between success rate and number of problems attempted. The significant parameter associated with path b indicated that an increase of number of problems attempted was related to a small increase in math performance.

The bootstrap test of Preacher and Hayes (2004) indicated that the indirect effect, similar to the difference between the total effect and the direct effect of success rate was estimated at −.45, and differed significantly from zero (bootstrapped 95%-confidence interval ranged from −1.50 to −.03). Hence, increasing success rate indirectly increased math improvement, because children in easier conditions attempted more problems in Math Garden than children in more difficult conditions. Difference scores were equal across grades, but

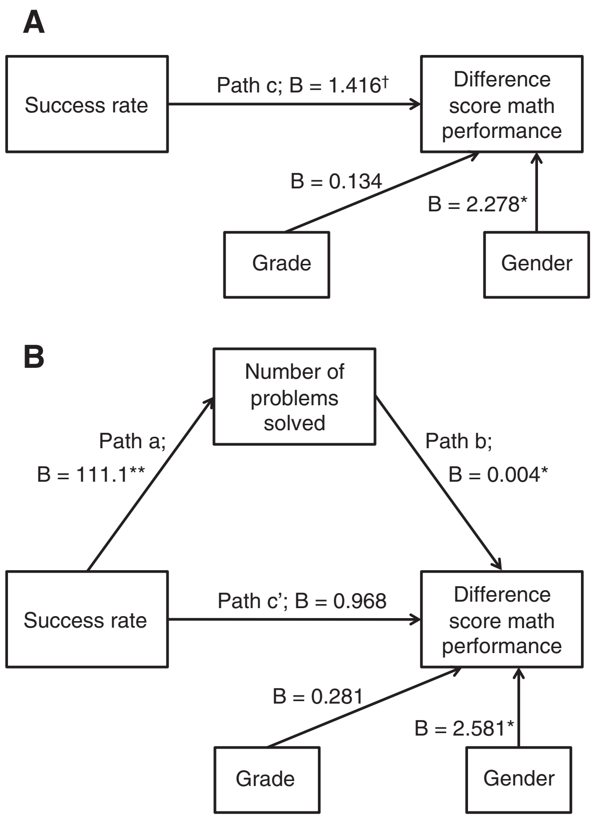

Fig. 3. Illustration of linear models for the effect of success rate on addition–subtraction difference scores. Fig. 3A shows the total effect whereas Fig. 3B shows the direct effect of success rate on addition–subtraction difference scores, controlling for the number of problems attempted in Math Garden. B: unstandardized coefficients; †$p<.1$; *$p<.05$; **$p<.01$.

differed for boys and girls. Girls' math performance improved more than that of boys.

3.6. Summary

Math anxiety scores lowered from pretest to posttest but the decrease was equal for children in the experimental conditions and children in the control condition. Generally, perceived math competence scores did not change from pretest to posttest. Addition–subtraction scores improved significantly from pretest to posttest for children in the experimental conditions but remained equal for children in the control condition. Improvements in math performance were most pronounced for the condition with the highest success rate and were higher for girls than for boys. Mediation analyses indicated that success rate indirectly affected improvement in math performance, through the number of problems attempted in Math Garden: the higher the success rate, the more problems were attempted, and the higher the improvement in math performance.

4. Discussion

In the present study, it was investigated whether practicing math with a computer-adaptive program would decrease math anxiety and increase perceived math competence because the program ensured that all children achieved high success rates in math. This was ensured because the computer-adaptive program (Math Garden; Klinkenberg et al., 2011) adjusted problem difficulty to the individual ability level. Given the reciprocal relationship between math anxiety and perceived competence on the one hand and math performance on the other hand (e.g., Hembree, 1990; Ma, 1999), we expected math performance to improve as well. Additionally, this study investigated whether level of success rate, manipulated in the computer-adaptive program, affected math anxiety, perceived math competence, and math performance. Children were in a control condition or in one of three experimental conditions, in which they used Math Garden for six weeks. Success rates in experimental conditions Difficult, Medium, and Easy were aimed at 60%, 75%, and 90%, respectively.

Math anxiety, "a person's negative affective reaction to situations involving numbers, math, and mathematics calculations" (Ashcraft & Moore, 2009, p. 197), has been found to correlate with math performance from as early as the second and third grades (Wu, Barth, Amin, Malcarne, & Menon, 2012). Although empirical grounds for an explanation of the onset of math anxiety are lacking (Ashcraft & Moore, 2009), many researchers agree that negative math experiences, like failures, may play an important role (e.g., Ashcraft & Krause, 2007; Ashcraft & Moore, 2009; Bekdemir, 2010; Rubinsten & Tannock, 2010). Hence, it was expected that leveling success rate, at a relatively high level, would reduce math anxiety. However, math anxiety scores improved equally for children in all conditions. This absence of effects of practicing with Math Garden on math anxiety may relate to the first assessment of the math anxiety questionnaire. Children seemed quite tense during this pretest, possibly caused by the announcement of the arrival of the experimenter, who was doing a research on math. On the posttest, math anxiety scores dropped greatly, also for children in the control condition. Questions referring to math tests contributed importantly to this drop. Timing of the posttest, shortly before summer holidays, may have played a role. Children may have realized that tests were less likely at this time of the year.

Perceived competence is related to constructs like self-efficacy and self-concept (Lee, 2009). In the present study, we asked children to compare their math competence to that of their peers. Although it is clear that evaluation of competence and achievement are related, the causal ordering is still undetermined (Guay et al., 2003). In addition, it is debated whether evaluation of competence and test anxiety are related or independent constructs (Bandalos et al., 1995; Lee, 2009). In the present study, we noted a difference between perceived math competence scores and math anxiety scores. Whereas math anxiety scores improved from pretest to posttest, perceived math competence scores only improved slightly from pretest to posttest, and only for children in the condition with a medium success rate. Results may have been modest because of the short period of practicing math with Math Garden.

There was a small but significant improvement of math performance for children who used Math Garden, but not for children following the regular curriculum. Moreover, improvement was highest for children in the condition with the highest success rate. The relation was mediated by the number of problems attempted in the computer-adaptive program. The higher the success rate, the more children played in Math Garden, and the larger the improvement in math performance. This result suggests that practicing math frequently at one's own ability level improves math performance, and that the experience of success stimulates this practice. Note, however, that even in the easy condition, children still solved 19% of the problems incorrectly, so the problems were not overly easy.

The finding that girls improved more than boys in math performance, after practicing with the computer-adaptive program, was unanticipated. Gender differences in improvement of math performance after training are scarcely reported (but see Timmermans, van Lieshout, & Verhoeven, 2007). The present results suggest that girls benefited more from tailor-made problem selection, high success rate, and making errors in private. It should be noted that we focused on two specific domains of math: addition and subtraction.

Males and females did not differ in levels of math anxiety and of perceived math competence. It was suggested in the introduction that gender differences in math anxiety might differ between cultures and may become more explicit with increasing age (see Hembree, 1990). A similar suggestion may hold for perceived math competence as the constructs are closely related (Bandalos et al., 1995; Lee, 2009). The absence of gender effects in math anxiety found in this study may be related to the general low level of math anxiety in The Netherlands (Lee, 2009).

Effects of practicing with Math Garden and manipulating success rate were observed for math performance but absent for math anxiety and perceived competence. The correlations between math anxiety and perceived math competence on the one hand and math performance on the other hand do support a relation between emotional experience of math and math performance. Several authors note that math anxiety may be developed because of negative math experiences in the classroom and suggest that mathematics anxiety is a phenomenon which begins at an early age (Ashcraft & Krause, 2007; Bekdemir, 2010; Newstead, 1998; Rubinsten & Tannock, 2010). Indeed, Newstead (1998) found effects of classroom practice reforms on math anxiety levels of nine- to eleven-year-olds. Hence, although math experiences with the computer-adaptive program in the present study may have been positive, the experiences may not outweigh (negative) math experiences in the past. An alternative explanation for the absence of effects concerning emotional experience of math is that this study was performed in a normal and not a clinical population.

The beneficial effects of practicing math with Math Garden, compared to a wait-list condition, might promote the use of computer-adaptive programs in teaching math. More research regarding the requirements of such a program is still desired. As instructions on the math curriculum were not given in the present study, it may be the case that the computer-adaptive practice was additional to the regular math lessons and that the beneficial effects were due to increased practice of math. Furthermore, any computer-adaptive program may work as they all adapt problem difficulty to the individual's ability level and provide privacy when making mistakes. It is, however, likely that other conditions are important as well. Math Garden includes many attractive features (e.g., high success rate, emphasis on the automation of basic math facts, immediate feedback, reward of both speed and accuracy, and colorful game-like environment) for children. The findings in the present study do support the hypothesis that high success rate is a crucial feature. The relation between increasing success rate and improving math performance, mediated by the number of problems attempted, demonstrates the importance for children of being successful in math.

References

Ashcraft, M. H., & Faust, M. W. (1994). Mathematics anxiety and mental arithmetic performance: An exploratory investigation. *Cognition & Emotion*, *8*, 97–125, http://dx.doi.org/10.1080/02699939408408931.

Ashcraft, M. H., & Krause, J. A. (2007). Working memory, math performance, and math anxiety. *Psychonomic Bulletin and Review*, *14*, 243–248, http://dx.doi.org/10.3758/BF03194059.

Ashcraft, M. H., & Moore, A. M. (2009). Mathematics anxiety and the affective drop in performance. *Journalf of Psychoeducational Assessment*, *27*(3), 197–205, http://dx.doi.org/10.1177/0734282908330580.

Bandalos, D. L., Yates, K., & Thorndike-Christ, T. (1995). Effects of math self-concept, perceived self-efficacy, and attributions for failure and success on test anxiety. *Journal of Educational Psychology*, *87*, 611–623, http://dx.doi.org/10.1037//0022-0663.87.4.611.

Bekdemir, M. (2010). The pre-service teachers' mathematics anxiety related to depth of negative experiences in mathematics classroom while they were students. *Educational Studies in Mathematics*, *75*, 311–328, http://dx.doi.org/10.1007/s10649-010-9260-7.

Bonnot, V., & Croizet, J. -C. (2007). Stereotype internalization and women's math performance: The role of interference in working memory. *Journal of Experimental Social Psychology*, *43*, 857–866, http://dx.doi.org/10.1016/j.jesp. 2006.10.006.

Chinn, S. (2009). Mathematics anxiety in secondary students in England. *Dyslexia*, *15*, 61–68, http://dx.doi.org/10.1002/dys.381.

Chiu, L. H., & Henry, L. L. (1990). Development and validation of the mathematics anxiety scale for children. *Measurement and Evaluation in Counseling and Development*, *23*, 121–127.

De Vos, T. (2010). *Manual Tempotoets Automatiseren*. Amsterdam: Boom Testuitgevers.

Guay, F., Marsh, H. W., & Boivin, M. (2003). Academic self-concept and academic achievement: Developmental perspectives on their causal ordering. *Journal of Educational Psychology*, *95*, 124–136, http://dx.doi.org/10.1037/0022-0663.95.1.124.

Harter, S. (1982). The perceived competence scale for children. *Child Development*, *53*, 87–97, http://dx.doi.org/10.1111/j.1467-8624.1982.tb01295.x.

Hembree, R. (1990). The nature, effects, and relief of mathematics anxiety. *Journal for Research in Mathematics Education*, *21*, 33–46, http://dx.doi.org/10.2307/749455.

Ho, H. -Z., Senturk, D., Lam, A. G., Zimmer, J. M., Hong, S., Okamoto, Y., et al. (2000). The affective and cognitive dimensions of math anxiety: A cross-national study. *Journal for Research in Mathematics Education*, *31*, 362–379, http://dx.doi.org/10.2307/749811.

Klinkenberg, S., Straatemeier, M., & van der Maas, H. L. J. (2011). Computer adaptive practice of maths ability using a new item response model for on the fly ability and difficulty estimation. *Computers in Education*, *57*, 1813–1824, http://dx.doi.org/10.1016/j.compedu.2011.02.003.

Krinzinger, H., Kaufmann, L., & Willmes, K. (2009). Math anxiety and math ability in early primary school years. *Journal of Psychoeducational Assessment*, *27*, 206–225, http://dx.doi.org/10.1177/0734282908330583.

Lee, J. (2009). Universals and specifics of math self-concept, math self-efficacy, and math anxiety across 41 PISA 2003 participating countries. *Learning and Individual Differences*, *19*, 355–365, http://dx.doi.org/10.1016/j.lindif.2008.10.009.

Liu, O. L., & Wilson, M. (2009). Gender differences in large-scale math assessments: PISA Trend 2000 and 2003. *Applied Measurement in Education*, *22*(2), 164–184.

Ma, X. (1999). A meta-analysis of the relationship between anxiety toward mathematics and achievement in mathematics. *Journal for Research in Mathematics Education*, *30*, 520–540, http://dx.doi.org/10.2307/749772.

Ma, X., & Xu, J. M. (2004). The causal ordering of mathematics anxiety and mathematics achievement: A longitudinal panel analysis. *Journal of Adolescence*, *27*, 165–179, http://dx.doi.org/10.1016/j.adolescence.2003.11.003.

Maris, G., & Van der Maas, H. L. J. (2012). Speed-accuracy response models: Scoring rules based on response time and accuracy. *Psychometrika*, http://dx.doi.org/10.1007/s11336-012-9288-y [Advance online publication].

Marsh, H. W., & Martin, A. J. (2011). Academic self-concept and academic achievement: Relations and causal ordering. *British Journal of Educational Psychology*, *81*, 59–77, http://dx.doi.org/10.1348/000709910X503501.

Newstead, K. (1998). Aspects of children's mathematics anxiety. *Educational Studies in Mathematics*, *36*, 53–71, http://dx.doi.org/10.1023/A:1003177809664.

Preacher, K. J., & Hayes, A. F. (2004). SPSS and SAS procedures for estimating indirect effects in simple mediation models. *Behavior Research Methods, Instruments, & Computers*, *36*(4), 717–731, http://dx.doi.org/10.3758/BF03206553.

Preacher, K. J., & Hayes, A. F. (2008). Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behavior Research Methods*, *40*, 879–891, http://dx.doi.org/10.3758/BRM.40.3.879.

Rosenthal, R., Rosnow, R. L., & Rubin, D. B. (2000). *Contrasts and effects sizes in behavioral research: A correlational approach.* Cambridge, UK: Cambridge Press.

Rubinsten, O., & Tannock, R. (2010). Mathematics anxiety in children with developmental dyscalculia. *Behavioral and Brain Functions*, *6*, 1–13, http://dx.doi.org/10.1186/1744-9081-6-46.

Timmermans, R. E., van Lieshout, E. C. D. M., & Verhoeven, L. (2007). Gender-related effects of contemporary math instruction for low performers on problem-solving behavior. *Learning and Instruction*, *17*(1), 42–54.

Valentine, J. C., DuBois, D. L., & Cooper, H. (2004). The relation between self-beliefs and academic achievement: A meta-analytic review. *Educational Psychologist*, *39*, 111–133, http://dx.doi.org/10.1207/s15326985ep3902_3.

Van der Maas, H. L. J., Dolan, C. V., Grasman, R. P. P. P., Wicherts, J. M., Huizenga, H. M., & Raijmakers, M. E. J. (2006). A dynamical model of general intelligence: The positive manifold of intelligence by mutualism. *Psychological Review*, *113*, 842–861, http://dx.doi.org/10.1037/0033-295X.113.4.842.

Van der Ven, S. H. G., Kroesbergen, E. H., Boom, J., & Leseman, P. P. M. (2012). The development of executive functions and early mathematics: A dynamic relationship. *British Journal of Educational Psychology*, *82*, 100–119, http://dx.doi.org/10.1111/j.2044-8279.2011.02035.x.

Veerman, J. W., Straathof, M. A. E., Treffers, Ph. D. A., Van den Bergh, B. R. H., & Ten Brink, L. T. (1997). *Dutch manual for perceived competence scale for children.* Amsterdam: Pearson.

Wu, S., Barth, M., Amin, H., Malcarne, V., & Menon, V. (2012). Math anxiety in second and third graders and its relation to mathematics achievement. *Frontiers in Psychology*, *3*(162), http://dx.doi.org/10.3389/fpsyg.2012.00162.