Numeracy
Advancing Education in Quantitative Literacy

Volume 7 | Issue 1 Article 4

2014

Sixth Graders Benefit from Educational Software when Learning about Fractions: A Controlled Classroom study

Susanne Scharnagl
University of Ulm, Germany, susanne.scharnagl@znl-ulm.de
Petra Evanschitzky
Judith Streb
Manfred Spitzer
Katrin Hille

Recommended Citation
Scharnagl, Susanne; Evanschitzky, Petra; Streb, Judith; Spitzer, Manfred; and Hille, Katrin (2014) "Sixth Graders Benefit from Educational Software when Learning about Fractions: A Controlled Classroom study," *Numeracy*: Vol. 7 : Iss. 1 , Article 4.
DOI: http://dx.doi.org/10.5038/1936-4660.7.1.4
Available at: http://scholarcommons.usf.edu/numeracy/vol7/iss1/art4

Sixth Graders Benefit from Educational Software when Learning about Fractions: A Controlled Classroom study

Abstract
This study analyses the effectiveness of an educational web-based software package for teaching mathematics in schools. In all, 864 sixth graders and their teachers took part in the controlled study. Students learned the addition and subtraction of fractions with (intervention group; n = 469) or without (control group; n = 395) the support of the educational software. Compared to the controls, students who used the software showed better results in the post-test. Gains were dose dependent and particularly marked in high-ability students and students with lower scores of math anxiety.

Keywords
media in education, improving classroom teaching, intelligent tutoring systems, interactive learning environments

Cover Page Footnote

Susanne Scharnagl is an educational scientist at the ZNL Transfer Centre for Neuroscience and Learning, University of Ulm, Germany. Her research interests include development of self-regulation in children and adults - especially of emotional self-regulation - and how to create a powerful student-teacher relationship.

Petra Evanschitzky is an educational scientist and project manager at the ZNL Transfer Centre for Neuroscience and Learning at the University of Ulm, Germany. Her work focusses on developing education concepts for kindergarten and primary school, supporting the professionals in educational institutions and managing research projects.

Judith Streb is a senior researcher at the ZNL Transfer Centre for Neuroscience and Learning, University of Ulm, Germany. She completed her Ph.D. in psychology. Her research interests include methodological issues, ambulatory assessment and the social and emotional development of children.

Manfred Spitzer is the Medical Director, Professor and Chairman of the Psychiatric Hospital at the University of Ulm, Germany. In 2004, he founded the ZNL Transfer Center for Neuroscience and Learning where he holds the position of Chairman. His research activities focus on higher cognitive functions at the interface between cognitive neuroscience, psychology, and psychopathology.

Katrin Hille is the Director of the ZNL Transfer Centre for Neuroscience and Learning at the University of Ulm, Germany. She completed her Ph.D. in psychology. As a psychologist she plans and directs evaluations of teaching and learning situations. Her research focusses especially on questions that arise from practitioners' perspectives.

Introduction

The use of computers in education provokes a wide range of opinions: Some say that without serious, systematic incorporation of information technology (IT), education systems will be unable to adapt to the demands of the labor market and future societies (Cabrol and Severin 2010). Others say that media have no more effect on learning than a grocery truck has on the nutritional value of the product it brings to market (Clark 1983) and notice that computers have come to classrooms but sit there mostly unused (Cuban 2001). Rather than being just useless, computers may even be an obstacle to teaching and learning as they not only cause distraction but also are a means to “outsource” mental activity, causing shallower levels of information processing and therefore decreased learning (Craik and Lockhart, 1972; Carr 2008; Ophir et al. 2009; Sparrow, Liu and Wegner 2011). In fact, small children have definitely been found to suffer from “educational” screen media software (Zimmerman et al. 2007; Lewin 2009).

In the light of this wide spectrum of opinions and given the fact that computers are distributed within the educational environments of an increasing number of countries by the millions,[1] it is of interest that sound data from empirical studies on the effects of various school programs hardly exist (Mervis 2004) and that empirical studies on the effects of specific IT hardware and software on mathematical achievement produced often negative results. In particular, a study funded by the U.S. Department of Education found that none of the four math software packages tested produced a statistically significant difference in achievement between control and treatment groups over the two-year study period (Dynarski et al., 2007; Campuzano et al. 2009; Mervis 2009).

However, meta-analyses suggest that the use of computer technology may have a positive effect on student’s achievement in formal face-to-face classrooms as compared to classrooms that do not use technology, with average effect sizes of a moderate level (d = 0.37, Hattie 2009; d = 0.35, Tamim et al. 2011). Interestingly, the reported average of effect sizes for educational math software was smaller (d = 0.21, Hattie 2009).

Meta-analyses further showed a considerable variance of the use of computers in education, i.e., the concepts and tools differ considerably between studies. Therefore, it is not only important to show that the use of

[1] One Laptop per Child (OLPC) sells laptops to governments in the developing world with the goal of distributing "one laptop per child". OLPC claims over 2.4 million laptops have been shipped. Peru so far has bought 860,000 laptops for schools; in Uruguay this number is 510,000; in Rwanda it is 110,000. http://laptop.org/map (accessed Sept 20, 2013).

computer technology has a positive effect in general but also to show the benefits of specific software in particular.

This study was set up to evaluate the educational math software product developed by bettermarks GmbH (Germany, Berlin), a company founded and especially dedicated to develop math training software for use in school settings.

Why target mathematics with software? An everyday understanding of mathematics makes it possible for a person to function in society. But although basic math skills are important for everyday life, many people report feeling anxious when required performing mathematically. These feelings may not come from the mathematics itself but rather form the way mathematics is presented in school (Geist 2010). Students' motivational and emotional attitudes toward mathematics are influenced greatly by students' perceptions of success. Given that math anxiety impedes mathematical thinking, a vicious circle has long been proposed (Butterworth 1999): Experiences of failure (in particular, public failure in the classroom) cause anxiety, which blocks creative thinking, which leads to less mathematical problem-solving skills, which cause additional experiences of failure. As computers are infinitely patient and forgiving, they may take away a considerable amount of negative emotions in math training that usually accompany interpersonal classroom-training of mathematically challenged students.

The software that is evaluated in this study covers the math curricula of Germany from form 4 through 10. The part covering fractions in form 6 is evaluated in this study. Fractions, in form of percentages and rates, are omnipresent in people's lives and important for everyday decision making (Tucker 2007). Yet, many students continue their education and become adults without ever understanding fractions, especially understanding that a fraction is a part of something (Tucker 2007; Schield 2007). Consider the following question on the TIMSS 8^{th}-grade test: Find the approximate value, to the closest integer, of the sum: 19/20 + 23/25. Possible answers were (a) 1, (b) 2, (c) 42, (d) 45 (quoted in Tucker 2007). While the correct answer is (b), the majority of U.S. students chose (c) or (d).

The educational math software that is evaluated in this study helps the students not just to learn fractions algorithmically, that is to understand and master the rules and steps in solving fraction problems; it helps them to develop an understanding of what fractions are. For example there are exercises with fractions in measurement (e.g., time, money, lengths) that helps students to understand that a fraction is a part of something.

In the light of the quantitative literacy movement, an evaluation of a software that is not just a simple drill-and-practice program but also one that promotes the teaching of mathematics in general and fraction in particular is of interest.

DOI: http://dx.doi.org/10.5038/1936-4660.7.1.4

Material and methods

Subjects and study design

The study was carried out in a pre-test–post-test, control-group design. Subjects were recruited through their schools. In the first step of recruitment, headmasters of schools were asked if they wanted to try new educational math software and take part in this study. If the headmasters agreed we asked the teachers who teach mathematics in 6th grade if they want to try the software and take part in this study with their students. In order to keep teacher- and school-related variance at a minimum, we tried to recruit at least two classes per school. Ideally, those two classes would have been taught by the same teacher, but this was only the case in seven sets of intervention and control classes. In the remaining 31 sets of intervention and control classes, we recruited the control classes by approaching the teacher of the parallel class. If the teachers of intervention and control classes agreed to take part in the study, the students and their parents were asked for participation.

Altogether 34 schools agreed to take part in the study. For the final data analysis, some of the schools, or classes, respectively, had to be excluded for various reasons as illustrated in Figure 1. Four schools dropped out because of technical problems (e.g., insufficient IT infrastructure as the software needed a fast connection to the Internet). Four more schools had to be excluded because the study was not carried out correctly. There were two teachers who—despite being part of the intervention group—did not use the system with their students. Another school was excluded because a teacher shortened the testing instrument as he considered it too difficult. The results of the fourth school were eliminated because of incorrect students' codes that prevented the comparison between pre and post achievement levels. Finally, eight schools were excluded because of a prolonged instruction that included a holiday period.

Data from 864 six grade students (421 male) were analyzed. Of these 469 students (225 male) were part of the intervention group, and 395 students (196 male) from the parallel classes of the same schools were part of the control group (Fig. 1).

The intervention

The intervention was based on the bettermarks online system. The curriculum for mathematics in grade 6 includes learning how to add and subtract fractions. The teachers and students of the control group tackled the topic as usual. Teachers and students in the intervention group were supported by the bettermarks online system during the time period when addition and subtraction of fractions was taught. Teachers used the bettermarks online system to generate, organize, and analyze exercise series for school and homework. These exercise series are suggested by the system and can be tuned by the teacher to tailor it to students' achievements and needs. The system automatically corrects the students' task's solutions and offers assistance to students if they made a mistake.

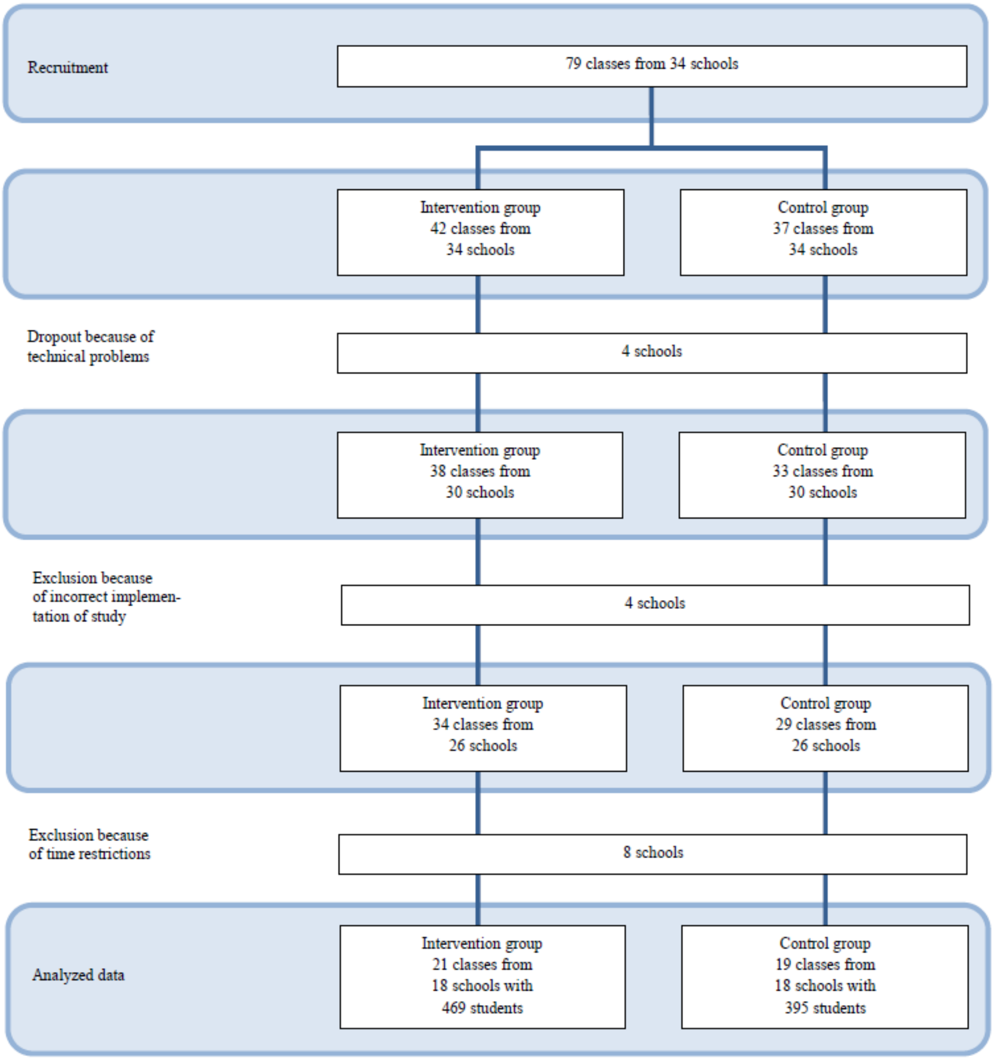

Figure 1: Flowchart with number of schools and classes from recruitment to data analysis

Every exercise is split up into each step necessary to solve the problem. The system guides the students step by step and gives them feedback after every step, if the solution was correct or not. Students get a second try in the case of a wrong submitted answer.

The bettermarks system goes further than just a simple drill-and-practice program (problem-solution-feedback). For example, after a wrong answer the student gets more than just a message "wrong," but also helpful tips on solving the problem (e.g., try simplifying more). Additionally, the system contains an online textbook where students can search for information regarding the problem (e.g., an explanation of what the common denominator is). This online textbook is linked directly to the exercises in such a way that students can navigate directly to the relevant chapter. In the case that this doesn't provide enough help and the students still fail on their second try, after the second wrong answer the student is provided with the solution, as well as the steps leading there and an explanation. Based on the level of competence, the students are offered further tasks on the topics that need practicing.

The teachers were asked to use the system as often as possible, in class and—at the least—for homework. The system supports teachers by providing practice tasks, tests, and homework. Additionally, the system allows the teacher to monitor the learning progress of individual students and of the entire class.

Variables

Mathematics achievement was assessed before and after the intervention with paper and pencil tests. The test before the intervention period addressed basic arithmetic competencies that are required to learn fractions, such as multiplication tables, division with remainder, combinations of addition and subtraction with multiplication, and lowest common multiples. The test after the intervention addressed fractional arithmetic—the topic taught during the intervention period. The tasks included addition, subtraction and comparison of fractions at various degrees of difficulty. Both, pre- and post-tests were available in two versions with identical tasks in different order to discourage cheating of students sitting next to each other. The tests were developed by Stein and Winter from the Institute of Education in Mathematics and Computer Science at the University of Münster. For data analysis, the percentage of correct answers was calculated from the pre- and post-test scores.

Students' level of math anxiety was measured with the German version of the index of mathematics anxiety that was used in PISA 2003 (Ramm et al. 2006). Students responded on a four-point scale to each of the following five statements: I often worry that it will be difficult for me in mathematics classes; I get very tense when I have to do mathematics homework; I get very nervous doing mathematics problems; I feel helpless when doing a mathematics problem; and I worry that I will get poor grades in mathematics. The level of agreement to the statements was summed so that higher scores indicate higher levels of mathematics anxiety.

The index of mathematics anxiety of 462 students (298 and 164 students in the intervention and the control condition, respectively) was assessed.

Multilevel modeling

Multilevel analyses assessed the effects of group, time, gender and math anxiety on the students' mathematical achievement. The present study was conducted in the school setting; therefore, student, classroom and school characteristics are confounded because students were not randomly assigned to groups. We used a three-level multilevel modeling approach to handle this data structure. The school level (level 3) was included to account for common variance, but no predictor was entered on this level. The classroom level (level 2) contained individual student characteristics such as gender and math anxiety. Level 2 also included information on group, i.e., intervention versus control. At level 1, the students' level, each student's successive measurements over time were defined.

Random-intercept models were estimated in which the intercepts on all three levels were allowed to vary randomly but with fixed effects for all predictor variables and cross-level interactions. The data were analyzed using the SPSS Mixed Procedure (SPSS 2002). The method of estimation applied for all models was restricted maximum likelihood. The following hypotheses about fixed effects were tested.

Hypothesis 1 states that the intervention works: students in the intervention group show higher gains in mathematical achievement than students in the control group. To test hypothesis 1, we examined the cross-level interaction between group (intervention vs. control) and time (pre vs. post).

Hypothesis 2 states that boys and girls respond differently to the intervention. To test hypothesis 2, we examined the cross-level interaction between group (intervention vs. control), time (pre vs. post) and gender (male vs. female).

Hypothesis 3 states that high and low math-achievers respond differently to the intervention. To test hypothesis 3, we examined the cross-level interaction between group (intervention vs. control), time (pre vs. post) and quartiles of pre-test math achievement (lower quartile split: 43.8; second quartile split: 54.2; upper quartile split: 63.3).

Hypothesis 4 states that students with high and low levels of math anxiety respond differently to the intervention. To test hypothesis 4, we examined the cross-level interaction between group (intervention vs. control), time (pre vs. post) and quartiles of math anxiety (lower quartile split: 7.75; second quartile split: 10.0; upper quartile split: 12.0). Students without data for math anxiety were excluded.

Hypothesis 5 states that intervention-group teachers who use the software more often than others have students who show higher gains in mathematical achievement than the students of the other teachers. Data showed a huge variation concerning the usage of the system by the teachers. One teacher offered his students only a single exercise series during the entire study period, whereas another teacher provided her students with 160 such exercise series. With a median split over the number of exercise series, the intervention teachers were categorized into "low usage" (less than 19 exercise series) and "high usage" (19 exercise series or more).

To test hypothesis 5, we looked only at the intervention group and examined the cross-level interaction between time (pre vs. post) and usage of the software.

DOI: http://dx.doi.org/10.5038/1936-4660.7.1.4

Results

Intraclass correlation (ICC)

The intraclass correlation was computed from the null model. The variance component output indicates that 35% of the math achievement variance occurred across schools and 9% across classrooms. The results suggest that the development of a multilevel model is warranted (Heck et al. 2010).

Does the intervention work?

Results for the hypothesis that the intervention works are shown in Table 1. We tested whether time and group as well as its interactions influenced the students' achievement. As indicated by the positive and significant interaction

Table 1.
Estimates of Fixed Effects for the hypothesis that the intervention works

Parameter	Estimate[a]	Std. Error	*df*	*t*-Score	Sig	CI 95% lower	CI 95% upper
Intercept	49.98	3.59	20.840	13.92	.000	42.51	57.45
[group=intervention]	0.69	2.21	26.499	0.31	.758	-3.85	5.22
[group=control]	0[b]	0	.	.	.	.	.
time	13.92	1.06	862	13.17	.000	11.85	15.99
[group=intervention] * time	3.06	1.43	862	2.13	.033	0.24	5.87
[group=control] * time	0[b]	0	.	.	.	.	.

a. Dependent Variable: percent.
b. This parameter is set to zero because it is redundant.

between group*time, the intervention group showed a significantly greater gain from pre-test to post-test compared to the control group. Students who learned fractions supported by the bettermarks online system achieved better results compared to their peers without the online system. While the pre-test results did not differ significantly, students of the intervention group achieved on average 3.06 points more in the post-test than the controls. The significant parameter time shows that the second test (post-test) was easier for the students than the first (pre-test).

Is the success of the intervention moderated by gender?

Results for the hypothesis that boys and girls respond differently to the intervention are shown in Table 2. We tested whether time, group, gender and its interactions influenced the students' achievement. The non-significant interaction indicated that we cannot say that girls and boys responded differently to the intervention. However generally, girls show on average 2.17 fewer points on math achievement.

Is the success of the intervention moderated by math achievement in pre-test?

Results for the hypothesis that high and low math-achievers respond differently to the intervention are shown in Table 3 and Figure 2. We tested

Table 2.
Estimates of Fixed Effects for the hypothesis that girls and boys respond differently to the intervention

Parameter	Estimate[a]	Std. Error	*df*	t-Score	Sig	CI 95% lower	CI 95% upper
Intercept	58.11	3.39	16.40	17.16	.000	50.95	65.27
gender [male=0]	-2.17	1.12	804.57	-2.21	.035	-4.22	-0.15
[group=intervention] * gender	-1.02	2.70	1678.45	-0.38	.705	-6.31	4.27
[group=control] * gender	0[b]	0	.	.	.	.	.
gender * time	2.52	2.64	862	0.96	0.34	-2.66	7.69
[group=intervention] * gender * time	-3.21	3.58	862	-0.90	0.37	-10.23	3.82
[group=control] * gender * time	0[b]	0	.	.	.	.	.

a. Dependent Variable: percent.
b. This parameter is set to zero because it is redundant.

Table 3.
Estimates of Fixed Effects for the hypothesis that high and low math achiever respond differently to the intervention

Parameter	Estimate[a]	Std. Error	*df*	*t*-Score	Sig	CI 95% lower	CI 95% upper
Intercept	67.97	2.62	37.72	25.96	.000	62.67	73.27
[pre-test_quartile=0_25]	-34.44	2.12	1637.56	-16.21	.000	-38.60	-30.27
[pre-test_quartile =25_50]	-20.85	2.27	1696.62	-9.19	.000	-25.30	-16.40
[pre-test_quartile =50_75]	-13.29	2.11	1691.77	-6.30	.000	-17.43	-9.15
[pre-test_quartile =75_100]	0[b]	0	.	.	.	.	.
[pre-test_quartile =0_25] * group[c]	3.80	2.39	173.83	1.59	.114	-0.92	8.52
[pre-test_quartile =25_50] * group	0.66	2.44	200.63	0.27	.788	-4.16	5.47
[pre-test_quartile =50_75] * group	0.76	2.38	176.97	0.32	.751	-3.93	5.45
[pre-test_quartile =75_100] * group	-3.37	2.36	162.54	-1.43	.155	-8.03	1.29
[pre-test_quartile =0_25] * time[d]	16.44	1.95	1674.46	8.45	.000	12.62	20.26
[pre-test_quartile =25_50] * time	16.06	2.40	1674.46	6.68	.000	11.34	20.77
[pre-test_quartile =50_75] * time	14.55	2.14	1674.46	6.80	.000	10.36	18.75
[pre-test_quartile =75_100] * time	9.10	2.03	1674.46	4.49	.000	5.13	13.08
[pre-test_quartile =0_25] * group * time	1.48	2.86	1674.46	0.52	.606	-4.13	7.09
[pre-test_quartile =25_50] * group * time	0.46	3.00	1674.46	0.15	.877	-5.42	6.34
[pre-test_quartile =50_75] * group * time	1.72	2.89	1674.46	0.60	.552	-3.94	7.38
[pre-test_quartile =75_100] * group * time	8.32	2.82	1674.46	2.96	.003	2.80	13.85

a. Dependent Variable: percent.
b. This parameter is set to zero because it is redundant.
c. Group: control = 0, intervention = 1
d. Time: pre = 0; post = 1

whether time, group, quartile of math achievement in pre-test and its interactions influenced the students' achievement.

As indicated by the positive and significant interaction between group * time * quartile, the students of the best quartile from the intervention group showed a significantly greater gain from pre-test to post-test; compared to the others, they gained on average 8.32 points more. Additionally, the interaction pre-test quartile * time shows what model one suggested already for the whole group: the post-test was easier than the pre-test for students of all quartiles.

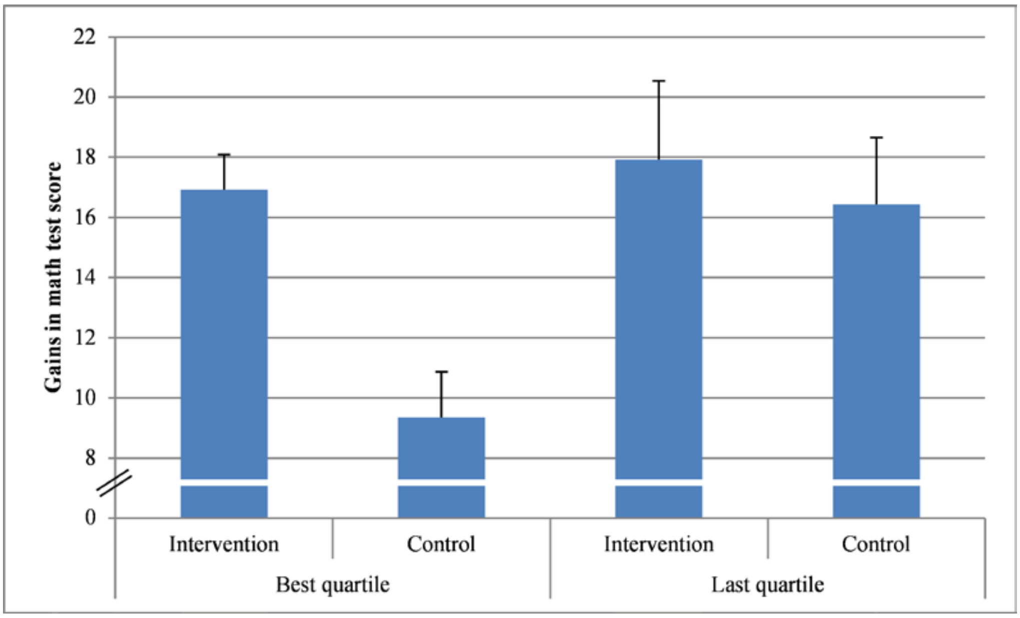

Figure 2: Gains in math test score (percentage correct on post-test minus percentage correct on pre-test) of students who belong to the last and to the best quartile regarding their results in the pre-test. Error bars represent SEM.

Is the success of the intervention moderated by math anxiety?

Results for the hypothesis that students with different levels of math-anxiety respond differently to the intervention are shown in Table 4. We tested whether time, group, quartile of math anxiety and its interactions influenced the students' achievement.

As indicated by the positive and significant interaction between group * time * quartile, the students with the lowest level of math anxiety showed a significantly greater gain from pre-test to post-test compared to the others. Students with the lowest level of math anxiety gained on average 9.42 points more. Additionally, the results show what model one suggested already for the whole.

Is the success of the intervention moderated by usage?

Results for the hypothesis that success of intervention is moderated by usage are shown in Table 5. We tested whether time and usage and their interaction influenced the students' achievement. This analysis was done for the intervention group only. As indicated by the positive and significant interaction between time * usage, the students with teachers who used the

Table 4.
Estimates of Fixed Effects for the hypothesis that students with high and low levels of math anxiety respond differently to the intervention

Parameter	Estimate[a]	Std. Error	*df*	*t*-Score	Sig	CI 95% lower	CI 95% upper
Intercept	40.76	5.06	37.17	8.05	.000	30.51	51.02
[math anxiety_quartile=0_25]	14.45	3.61	821.98	4.00	.000	7.37	21.54
[math anxiety_quartile =25_50]	5.84	3.38	820.62	1.73	.084	-0.79	12.47
[math anxiety_quartile =50_75]	5.44	4.27	826.27	1.27	.204	-2.95	13.83
[math anxiety_quartile =75_100]	0[b]	0	.	.	.	.	.
[math anxiety_quartile =0_25] * group[c]	-3.13	4.41	39.71	-0.71	.481	-12.05	5.78
[math anxiety_quartile =25_50] * group	2.23	4.04	28.36	0.55	.585	-6.04	10.50
[math anxiety_quartile =50_75] * group	2.59	4.97	64.30	0.52	.604	-7.34	12.53
[math anxiety_quartile =75_100] * group	7.79	4.65	49.15	1.67	.100	-1.56	17.14
[math anxiety_quartile =0_25] * time[d]	13.28	2.99	454	4.45	.000	7.41	19.15
[math anxiety_quartile =25_50] * time	18.42	2.49	454	7.41	.000	13.53	23.30
[math anxiety_quartile =50_75] * time	18.68	4.08	454	4.58	.000	10.66	26.71
[math anxiety_quartile =75_100] * time	17.54	3.26	454	5.38	.000	11.13	23.96
[math anxiety_quartile =0_25] * group * time	9.42	3.77	454	2.50	.013	2.01	16.84
[math anxiety_quartile =25_50] * group * time	2.01	3.13	454	0.64	.521	-4.14	8.15
[math anxiety_quartile =50_75] * group * time	2.51	4.74	454	0.53	.596	-6.80	11.83
[math anxiety_quartile =75_100] * group * time	-7.03	4.23	454	-1.66	.097	-15.35	1.28

a. Dependent Variable: percent.
b. This parameter is set to zero because it is redundant.
c. Group: control = 0, intervention = 1
d. Time: pre = 0; post = 1

Table 5.
Estimates of Fixed Effects for the hypothesis that students respond differently to the intervention in accordance with the usage of the system through their teachers

Parameter	Estimate[a]	Std. Error	df	t-Score	Sig	CI 95% lower	CI 95% upper
Intercept	59.35	3.82	11.07	15.53	.000	50.94	67.75
Usage[b]	0.27	6.68	10.48	0.04	.969	-14.53	15.07
Usage * time	10.49	2.48	880.41	4.24	.000	5.63	15.36

a. Dependent Variable: percent.
b. Usage: low = 0; high = 1

software more often showed a significantly greater gain from pre-test to post-test compared to the students of the teachers who used the online system sparingly. Students with more usage gained on average 10.49 points more. Additionally, the non-significant effect of usage shows that the difference between these groups was not present at pre-test.

DOI: http://dx.doi.org/10.5038/1936-4660.7.1.4

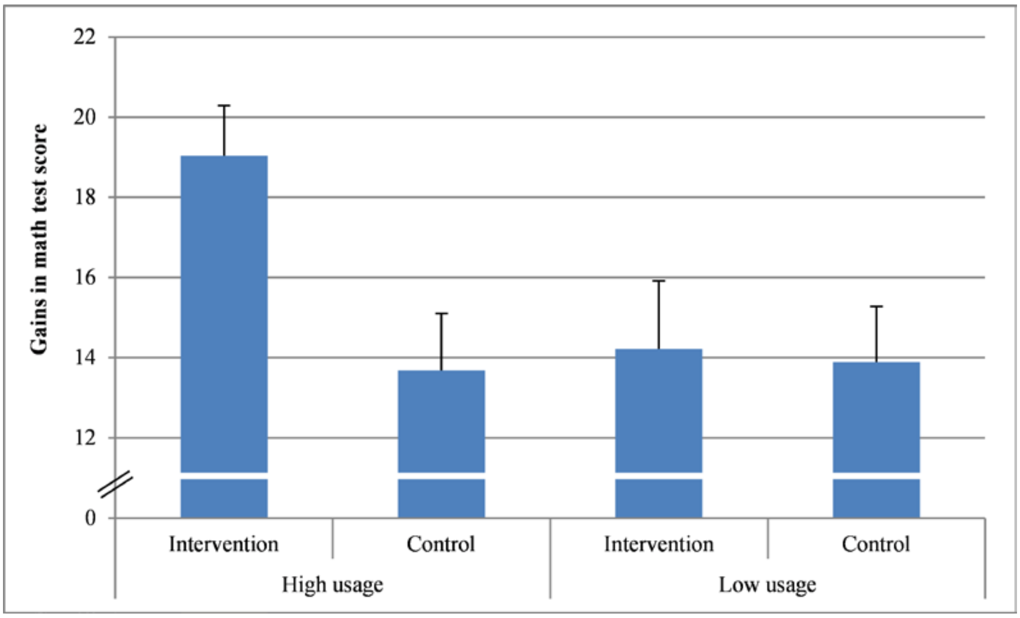

Figure 3: Gains in math test score (post percent – pre percent) of students with teachers who deployed more and less exercise series. Error bars represent SEM.

Discussion and Conclusions

In this study, we demonstrate that students learning to add and subtract fractions benefit from the use of an educational Web-based software package for teaching mathematics in schools (bettermarks online system). Students in the intervention group achieved on average 3.06 points more in the post-test than students in the controls. We cannot say that the effect of the intervention was larger in male students compared to female students, even though this is a result commonly found in studies on computer-assisted instruction (Hattie 2009). However, the general math achievement of male students in our sample was on average 2.17 points higher than that of female students. Such a result was also found in the Programme for International Student Assessment (PISA) in 2009 where boys outperformed girls in mathematics in Germany (*d*=.16) and across OECD nations (*d*=.13) (Reilly 2012).

One factor that moderated the effect of intervention was students' level of math anxiety. The intervention worked best for students with low math anxiety. Students in the quartile with the least math anxiety gained on average 9.42 points more in the post test.

A second factor that moderated the effect of intervention was students' level of math achievement in the pre-test. The students of the best quartile from the intervention group showed a significantly greater gain from pre-test to post-test compared to the best quartile from the controls. Students of the other quartiles from the intervention group showed also numerically greater gain compared to the controls, but the effects failed to reach the level of significance.

The lack of effect in students with low level of math achievement or with high level of math anxiety needs further exploration, as it may be due to several reasons. Firstly, the lack of basic arithmetic knowledge hinders the acquisition of knowledge on how to add and subtract fractions. Although the intervention system is programmed to remedy knowledge gaps by offering easier tasks, this option was not offered to the students in this study. A second reason for the lack of effect may lie in the short duration of the intervention. Possibly, the time allocated to grasp the concept of fractions was too short for the group of students with high levels of math anxiety to build up a substantial understanding, regardless of the amount and type of support.

A third factor that moderates the effect of the intervention is the usage of the system by the teachers. Students whose teachers deployed more exercise series benefitted more than the students whose teachers deployed only a few. In other words, the study uncovered a dose-dependent effect: the more the system was used, the larger the gains. Of course, one might speculate that, in general, committed teachers are better teachers, and in any case, teachers play an important role when estimating the benefit of educational software. However, the students of the two different teacher groups did not differ in the math achievement prior to intervention.

Qualitative data suggested that students were fond of the educational software. They answered open questions and responded that they especially enjoyed the immediate feedback provided by the system. Reimer and Moyer (2005) also report that users liked the feedback features of the computer systems they evaluated. Hattie (2009) highlights the important role of feedback when he summarizes in his meta-analysis that the use of computers is more effective when feedback is optimized. Our subjects considered the feedback of the system very helpful, especially for homework. Students get feedback for traditional homework only at the next lesson at school, if at all. In contrast, the system gives immediate and specific feedback on entering the result while students do their homework. The teachers also liked the system especially for homework. Using the system, teachers didn’t have to review the homework with students and can use the time for teaching.

In addition, the explanations provided by the system were highly valued by the users. The students were able to access information that helped them to understand concepts and to solve tasks on demand. They recommended to the software developer to expand this function.

Finally, although a positive effect of the bettermarks mathematics online training software on student achievement in formal face-to-face classrooms was shown, the study cannot clarify the reason for the positive effect. The reason could be direct, i.e., the educational software provides support for learning through individualized training. The reason could also be indirect in that the educational software encourages the students to spend more time on their math homework as they enjoy it more or find themselves able to do it. Lastly, the reason could be a combination of both a direct and an indirect effect. While it is interesting from a theoretical point of view to disentangle the effects, from a practical point of view, the results support the practicability of this kind of educational software for students and teachers who want to improve math achievement.

References

Butterworth, Brian. 1999. *What counts: How every brain is hardwired for math*. New York: The Free Press.

Cabrol, Marcelo, and Eugenio Severin. 2010. ICT to improve quality in education – A conceptual framework and indicators in the use of information communication technology for education (ICT4E). In *Assessing the effects of ICT in education: Indicators, criteria and benchmarks for international comparisons*, ed. Friedrich Scheuermann and Francesc Pedró, 83–106. Luxemburg: European Union/OECD.

Campuzano, Larissa, Mark Dynarski, Roberto Agodini, and Kristina Rall. 2009. *Effectiveness of reading and mathematics software products: Findings from two Student cohorts.* NCEE 2009-4041. Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.

Carr, Nicholas. 2008. Is Google making us stupid? Why you can't read the way you used to. *The Atlantic Monthly* 302(1): 56–63.

Clark, Richard E. 1983. Reconsidering research on learning from media. *Review of Educational Research* 53(4): 445–459. http://dx.doi.org/10.3102/00346543053004445

Craik, Fergus I M, and Robert S Lockhart. 1972. Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior* 11(6): 671–684. http://dx.doi.org/10.1016/S0022-5371(72)80001-X

Cuban, Larry. 2001. *Oversold and underused: computers in the classroom*. Cambridge, MA: Harvard University Press.

Dynarski, Mark, Roberto Agodini, Sheila Heaviside, Timothy Novak, Nancy Carey, Larissa Campuzano, Barbara Means, Robert Murphy, William Penuel, Hal Javitz, Deborah Emery, and Willow Sussex. 2007. *Effectiveness of reading and mathematics software products: Findings from the first student cohort*. Washington, D.C: U.S. Department of Education, Institute of Education Sciences.

Hattie, John A.C. 2009. *Visible learning: A synthesis of meta-analyses relating to achievement*. New York: Routledge.

Heck, R. H., S. L. Thomas, and L. N. Tabata. 2010. *Multilevel and longitudinal modeling with IBM SPSS*. New York: Taylor & Francis.

Lewin, Tamar. 2009. No Einstein in your crib? Get a refund. *The New York Times*, September 24, A1.

Mervis, Jeffrey. 2004. Meager evaluations make it hard to find out what works. *Science* 304(5677): 1583. http://dx.doi.org/10.1126/science.304.5677.1583

———. 2009. Study questions value of school software for students. *Science* 323(5919): 1277. http://dx.doi.org/10.1126/science.323.5919.1277

Ophir, Eyal, Clifford Nass, and Anthony D Wagner. 2009. Cognitive control in media multitaskers. *Proceedings of the National Academy of Sciences of the United States of America* 106(37): 15583–15587. http://dx.doi.org/10.1073/pnas.0903620106

Ramm, Gesa, Manfred Prenzel, and Jürgen Baumert. 2006. *PISA 2003. Dokumentation der Erhebungsinstrumente*. Münster: Waxmann.

Reilly, David. 2012. Gender, culture, and sex-typed cognitive abilities. *PloS one* 7(7): 39904. http://dx.doi.org/10.1371/journal.pone.0039904

Reimer, Kelly, and Patricia S Moyer. 2005. Third-graders learn about fractions using virtual manipulatives: A classroom study. *Journal of Computers in Mathematics and Science Teaching* 24(1): 5–25.

Schield, Milo 2007. Quantitative literacy and school mathematics: Percentages and fractions. In *Calculation vs. context: Quantitative literacy and its implications for teacher education*, ed. Bernard L. Madison and Lynn A. Steen, 87–107. Racine: Johnson Foundation.

Sparrow, Betsy, Jenny Liu, and Daniel M Wegner. 2011. Google effects on memory: Cognitive consequences of having Information at our fingertips. *Science* 333(6043): 776–778. http://dx.doi.org/10.1126/science.1207745

SPSS. 2002. *Linear mixed effects modeling in SPSS: An introduction to the mixed procedure (Technical report LMEMWP-1002)*. Chicago: Author.

Tamim, Rana M, Robert M Bernard, Eugene Borokhovski, Philip C Abrami, and Richard F Schmid. 2011. What forty years of research says about the impact of technology on learning: A second-order meta-analysis and validation study. *Review of Educational Research* 81(1): 4–28. http://dx.doi.org/10.3102/0034654310393361

Tucker, Alan. 2007. Fractions and units in everyday life. In *Calculation vs. context: Quantitative literacy and its implications for teacher education*, ed. Bernard L. Madison and Lynn A. Steen, 75–86. Racine: Johnson Foundation.

Zimmerman, Frederick J, Dimitri A Christakis, and Andrew N Meltzoff. 2007. Associations between media viewing and language development in children under age 2 years. *The Journal of Pediatrics* 151(4): 364–368. http://dx.doi.org/10.1016/j.jpeds.2007.04.071

DOI: http://dx.doi.org/10.5038/1936-4660.7.1.4