Calcularis Improvements in Percentiles

Christian Voegeli, cvoegeli@dybuster.com, Nov. 2018

Summary

This document contains the percentile improvements found in the Calcularis studies published as [Kaes13], [Aster15] and [Rauscher16]. They were not reported in the publications themselves, which focused on measuring raw values, because raw values tend to measure the progress for children with learning disabilities more transparently than percentile scores.

The test used in Calcularis studies as a pre- and post-test, for which standardized percentiles are available, is the test called "HRT" or "**H**eidelberger **R**echen**t**est". The percentile improvements for this test are reported here.

Table of Contents

1 Training Effects in Percentiles for Study of [Kaes2013]

1.1 Summary of Raw Value Results

The raw values for HRT Addition and HRT Subtraction (see [Kaes2013]) showed significant within-group improvements both for addition at $p < 0.05$ as well as subtraction at $p < 0.001$. They also showed that improvement in addition occurred with a smaller effect size than improvement in subtraction. Hence the between-group difference between training group and waiting group was significant for subtraction at $p < 0.01$, but not significant for addition. Due to the user adaptation of Calcularis, which had the participants train more subtraction than addition, this difference in addition and subtraction had to be expected. The same structure of the results can also be seen in the percentile improvements.

1.2 Improvements in Percentile Ranks

The largest improvement in percentile ranks was achieved by the **training group (TG) in subtraction** tasks. They rose from an average **rank of 12.9 to rank 29.0** from t_1 to t_2 i.e. within **6 weeks** of training, which resulted in a significant within-group i.e. t-score improvement. At the same time, the waiting group (WG) hardly improved. Hence also the between-group improvement was significant.

With respect to **addition**, the training group improved from average **rank 13.9 to 21.2** from t_1 to t_2, which as such is also very positive and significant as a within-group improvement. The waiting group improved from rank 14.6 to 17.4 in the same 6 weeks. The waiting group did not improve significantly. However, due to the change, the between-group difference for addition is not significant. As explained in [Kaes13], the individualization of the user adaptation of Calcularis leads to subtraction being practiced more often and hence to larger effect sizes for subtraction.

For the sake of completeness, also the percentile ranks at t_3 are shown in Table 1. From t_2 to t_3, both groups worked with Calcularis, so only within-group improvements can be measured. At the same time, the HRT is a test with quarterly norms, and the norms changed between t_2 and t_3. Therefore also within-group comparisons between t_2

and t_3 are of limited information, especially because the difference from t_2 to t_3 is 6 weeks, whereas the timeframe difference for the norms at t_2 and t_3 is 3 months.

There is however still a very positive interpretation for Calcularis: Without any intervention, one would have to expect subjects to get lower percentile ranks after 6 weeks, if the norm timeframe is 3 months. The subjects in the studies however managed to improve their percentile ranks from t_2 to $t_3$despite these different norms and timeframes, or at least did not lose ground.

		t_1	t_2	t-score	F-score[b]	t_3	t-score (t_2-t_3)
HRT Addition[a]	*TG*	13.9 (21.7)	21.2 (23.6)	-2.36*	1.18	23.3 (21.8)	-0.74
	WG	14.6 (13.6)	17.4 (15.6)	-1.03		21.1 (23.7)	-1.05
HRT Subtraction[a]	*TG*	12.9 (19.8)	29.0 (27.5)	-3.49**	6.19*	28.5 (28.3)	0.15
	WG	20.3 (15.6)	20.8 (16.0)	-0.10		22.8 (19.8)	-0.52

Table 1: Training effects of training group TG (n = 15) and waiting group WG (n = 17) on mathematical performance measured in percentile ranks (Means (SD)).

[+] p <.1,* p < .05,** p < .01,*** p < .001
[a] percentile ranks
[b] time (t_1–t_2) x group

In general, the improvements measured in percentile ranks are slightly less significant than their corresponding raw values. E.g. the t-score (t_1-t_2) for subtraction is significant at p < .001 for the raw values and "only" at p < .01 for percentiles. This is most probably a direct cause of the inherently weaker discrimination of improvements of under-performers by percentile scores as opposed to raw value scores.

2 Training Effects for Study of [Aster2015] and [Rauscher2016]

2.1 Summary of Study Design

In this study, there were a total of 6 groups: Half of the subjects were children with developmental dyscalculia (DD), and half of the subjects were normally achieving children (CC). Both were split up into a training group (CAL), a waiting group (WG) and a spelling training group (ST). This spelling training group performed a computerized spelling training to analyze possible effects of training with a computer as such i.e. in a different domain:

	Training Group	**Waiting Group**	**Spelling Training**
DD	CAL	WG	ST
CC	CAL	WG	ST

Table 2: Organization of the 6 groups in the study.

For the dyscalculic children (DD), there were no significant differences between the three groups for gender, age, math performance or control variables (intelligence, writing, reading) in the initial diagnostic procedure. Also amongst the three groups of the normally achieving children (CC), there were no significant differences.

The children trained for a total of 30 sessions of 20 minutes each. Children managing to train 5 times a week hence finished the training period in 6 weeks. Children, which worked less regularly, were allowed to finish the 30 sessions within at most 8 weeks. Children which did not practice 30 times in 8 weeks were excluded from the results.

2.2 Improvements in Percentile Ranks for Children with Developmental Dyslexia

As in the study summarized in Section 1, the training group (CAL) improved with regard to mathematical performance: The group x time interaction was significant regarding addition with medium effect size and regarding subtraction with large effect size. Furthermore, between-group comparison (CAL vs. WG, CAL vs. ST)

revealed **significant group x time interactions regarding subtraction**. With regard to addition, a significant group x time interaction was found for the comparison of the training group (CAL) with spelling training group (ST), but not for the comparison of the training group (CAL) with the waiting group (WG). Again, this difference may be explained by the individual user adaptation, which lead to practicing more subtraction than addition.

Looking at the percentile rank changes in absolute values reveals a large improvement for the CAL groups. For addition, they improved from an average rank of 4.46 to 11.30, and for subtraction, they improved from 4.36 to 9.74. Furthermore, comparison of the Calcularis training group (CAL) and the spelling training group (ST) also show significant group x time interactions. This proves that **the improvements achieved by Calcularis are caused by the domain specific training** provided by Calcularis and not by possible general effects of computer training.

Outcome Parameter	Group	n	t1	t2		*F*	*p*	η2
			M (SD)	*M (SD)*				
HRT (addition)[a]	CAL	23	4.46 (4.11)	11.30 (12.59)	overall	3.63	.032	.102
	WG	22	9.10 (10.36)	12.28 (17.57)	CAL-WG	1.34	.254	.030
	ST	22	10.51 (12.92)	8.91 (9.26)	CAL-ST	7.23	.010	.144
HRT (subtraction)[a]	CAL	23	4.36 (4.83)	9.74 (8.29)	overall	7.09	.002	.181
	WG	22	7.79 (11.50)	6.97 (10.43)	CAL-WG	11.15	.002	.206
	ST	22	5.78 (5.15)	6.61 (5.34)	CAL-ST	6.61	.014	.133

[a] percentile rank

Table 3: Training effects of the Calcularis group (CAL), waiting group (WG) and spelling training group (ST) of the children with mathematical learning disabilities (DD).

2.3 Improvements in Percentile Ranks for Children without Difficulties in Math

Also for children without difficulties in math, improvements in subtraction are larger than improvements in addition: **For addition, they increased from rank 47.2 to 59.7, and for subtraction, they improved from 48.1 to 64.6**. With regard to addition no significant group x time interaction was found. With regard to subtraction, study results demonstrated a significant group x time interaction with medium effect size. Further analyses demonstrated that children of the training group demonstrated a higher benefit than the spelling training group.

The interaction between group x time was not significant for the comparison of the Calcularis and waiting group.

Outcome Parameter	Group	n	t1	t2		F	p	η2
			M (SD)	M (SD)				
HRT (addition)[a]	CAL	20	47.20 (26.30)	59.70 (24.02)	overall	1.39	.256	.043
	WG	21	42.90 (30.51)	59.57 (24.53)	CAL-WG	0.40	.529	.010
	ST	24	38.29 (23.06)	44.75 (25.13)	CAL-ST	0.88	.353	.021
HRT (subtraction)[a]	CAL	20	48.10 (25.18)	64.60 (22.18)	overall	3.51	.036	.102
	WG	21	47.95 (26.04)	55.14 (27.87)	CAL-WG	2.66	.111	.064
	ST	24	43.17 (24.10)	45.83 (28.63)	CAL-ST	6.75	.013	.139

[a] percentile rank

Table 4: Training effects of the Calcularis group (CAL), waiting group (WG) and spelling training group (ST) of the children without difficulties in math (CC).

In this data set for the children without difficulties in math, the improvement for the waiting group (WG) is unexpectedly high compared especially to the improvement of the spelling training group (ST). The cause for this improvement could further be investigated. Independent of this cause, however, the comparison of the DD and CC groups leads to the conclusion that **children with DD profit more from the Calcularis training than the children without difficulties**. The absolute improvement in percentiles is larger for children without difficulties, but it is also larger for the control groups WG and ST without difficulties. This larger absolute improvement in percentiles might

be caused by measuring improvements around the mean, where smaller changes in raw values result in larger changes in percentile scores, because there are more children around the mean of a normal distribution then at its tails. Another cause could be that students without DD profited more from school instruction.

3 Interpretation

The improvements measured in percentile ranks are very promising for Calcularis: In both studies, they showed a **significant improvement for subtraction with large effect sizes** after only 6 weeks of training between the training and control groups for children with dyscalculia, despite the general disadvantages of measuring improvements in percentile ranks for children with learning difficulties. The improvements in addition were less significant than improvements in subtraction, because of the individual user adaptation: The study participants started at a higher level of correctly solved tasks for addition than for subtraction, and hence Calcularis by its adaptive nature asked more subtraction than addition tasks. Nevertheless, within-group improvements (group x time interactions) for addition are also significant.

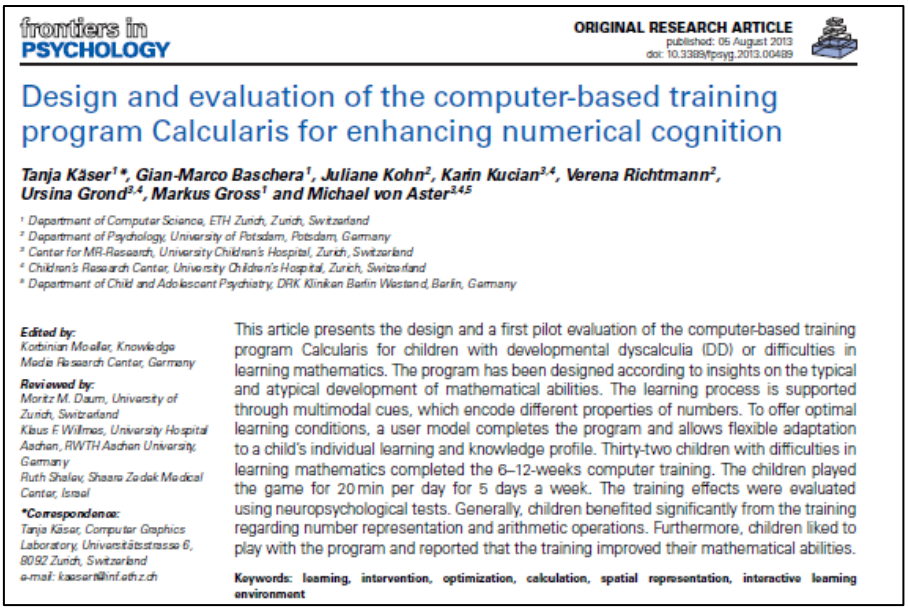
Over both studies, the **average net increase in percentile ranks over a 6 week training period is 6 ranks** for subtraction (weighted average improvement of training groups (TG/CAL): 6.34; weighted average improvement of control groups (WG/WG/ST): 0.35). As the training group in [Kaes2013] showed, this improvement is observable for at least 12 weeks, i.e. for a **12 week training, a net increase of 12 percentile ranks can be expected**, also across different norm time frames. From our point of view, that is a **tremendous increase** in such a short period of time, especially taking into account the little effort it required from educators (almost no effort) and the little progress the children made otherwise.

We are aware that interpolations to larger time spans have not been shown, and that assuming a steady improvement at one rank per week for e.g. half a year is just a hypothesis. Nevertheless, we think any parent or educator would be glad to take a child from e.g. rank 10 to rank 35 (and thus most likely to passing grades) within half a year.

4 References

[Kaes2013]

T. Kaseser, G.-M. Baschera, J. Kohn, K. Kucian, V. Richtmann, U. Grond, M. Gross, and M. von Aster. **Design and evaluation of the computer-based training program Calcularis for enhancing numerical cognition**, *Frontiers in Developmental Psychology*, 4: 489, 2013



frontiers in PSYCHOLOGY

ORIGINAL RESEARCH ARTICLE
published: 05 August 2013
doi: 10.3389/fpsyg.2013.00489

Design and evaluation of the computer-based training program Calcularis for enhancing numerical cognition

Tanja Käser[1]*, Gian-Marco Baschera[1], Juliane Kohn[2], Karin Kucian[3,4], Verena Richtmann[2], Ursina Grond[3,4], Markus Gross[1] and Michael von Aster[3,4,5]

[1] Department of Computer Science, ETH Zurich, Zurich, Switzerland
[2] Department of Psychology, University of Potsdam, Potsdam, Germany
[3] Center for MR-Research, University Children's Hospital, Zurich, Switzerland
[4] Children's Research Center, University Children's Hospital, Zurich, Switzerland
[5] Department of Child and Adolescent Psychiatry, DRK Kliniken Berlin Westend, Berlin, Germany

Edited by:
Korbinian Moeller, Knowledge Media Research Center, Germany

Reviewed by:
Moritz M. Daum, University of Zurich, Switzerland
Klaus F. Willmes, University Hospital Aachen, RWTH Aachen University, Germany
Ruth Shalev, Shaare Zedek Medical Center, Israel

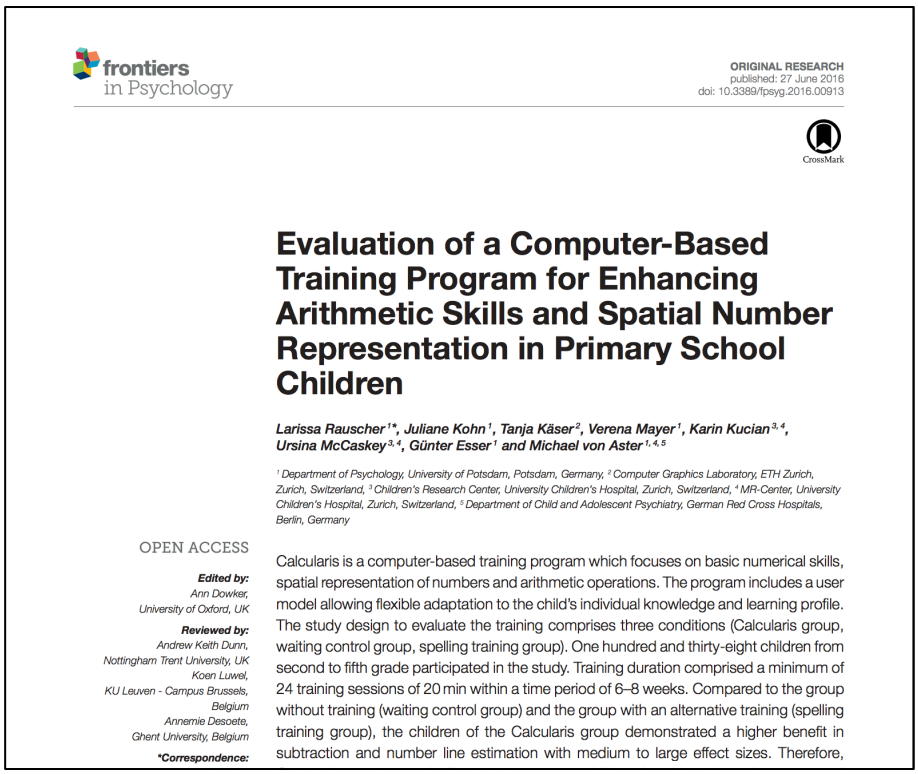
*Correspondence:
Tanja Käser, Computer Graphics Laboratory, Universitätsstrasse 6, 8092 Zurich, Switzerland
e-mail: kaeser@inf.ethz.ch

This article presents the design and a first pilot evaluation of the computer-based training program Calcularis for children with developmental dyscalculia (DD) or difficulties in learning mathematics. The program has been designed according to insights on the typical and atypical development of mathematical abilities. The learning process is supported through multimodal cues, which encode different properties of numbers. To offer optimal learning conditions, a user model completes the program and allows flexible adaptation to a child's individual learning and knowledge profile. Thirty-two children with difficulties in learning mathematics completed the 6–12-weeks computer training. The children played the game for 20 min per day for 5 days a week. The training effects were evaluated using neuropsychological tests. Generally, children benefited significantly from the training regarding number representation and arithmetic operations. Furthermore, children liked to play with the program and reported that the training improved their mathematical abilities.

Keywords: learning, intervention, optimization, calculation, spatial representation, interactive learning environment

[Rauscher2016]

L. Rauscher, J. Kohn, T. Käser, V. Mayer, K. Kucian, U. McCaskey, G. Esser, and M. von Aster. **Evaluation of a Computer-Based Training Program for Enhancing Arithmetic Skills and Spatial Number Representation in Primary School Children**, Frontiers in Developmental Psychology, 7:913, 2016



frontiers in Psychology

ORIGINAL RESEARCH
published: 27 June 2016
doi: 10.3389/fpsyg.2016.00913

Evaluation of a Computer-Based Training Program for Enhancing Arithmetic Skills and Spatial Number Representation in Primary School Children

Larissa Rauscher[1]*, Juliane Kohn[1], Tanja Käser[2], Verena Mayer[1], Karin Kucian[3,4], Ursina McCaskey[3,4], Günter Esser[1] and Michael von Aster[1,4,5]

[1] Department of Psychology, University of Potsdam, Potsdam, Germany, [2] Computer Graphics Laboratory, ETH Zurich, Zurich, Switzerland, [3] Children's Research Center, University Children's Hospital, Zurich, Switzerland, [4] MR-Center, University Children's Hospital, Zurich, Switzerland, [5] Department of Child and Adolescent Psychiatry, German Red Cross Hospitals, Berlin, Germany

OPEN ACCESS

Edited by:
Ann Dowker, University of Oxford, UK

Reviewed by:
Andrew Keith Dunn, Nottingham Trent University, UK
Koen Luwel, KU Leuven - Campus Brussels, Belgium
Annemie Desoete, Ghent University, Belgium

*Correspondence:

Calcularis is a computer-based training program which focuses on basic numerical skills, spatial representation of numbers and arithmetic operations. The program includes a user model allowing flexible adaptation to the child's individual knowledge and learning profile. The study design to evaluate the training comprises three conditions (Calcularis group, waiting control group, spelling training group). One hundred and thirty-eight children from second to fifth grade participated in the study. Training duration comprised a minimum of 24 training sessions of 20 min within a time period of 6–8 weeks. Compared to the group without training (waiting control group) and the group with an alternative training (spelling training group), the children of the Calcularis group demonstrated a higher benefit in subtraction and number line estimation with medium to large effect sizes. Therefore,

[Aster2015]

M. von Aster, L. Rauscher, K. Kucian, T. Kaeser, U. McCaskey, and J. Kohn. **Calcularis: Evaluation of a computer-based learning program for enhancing numerical cognition for children with developmental dyscalculia**, AACAP poster, 2015

Calcularis: Evaluation of a computer-based learning program for enhancing numerical cognition for children with developmental dyscalculia

Michael von Aster[a,c,e], Larissa Rauscher[a], Karin Kucian[c,d], Tanja Käser[b], Ursina McCaskey[c,d], Juliane Kohn[a]

a Department of Psychology, University of Potsdam, Karl-Liebknecht-Str. 24-25, 14476 Potsdam, Germany
b Computer Graphics Laboratory, ETH Zurich, Universitätsstrasse 6, 8092 Zurich, Switzerland
c MR-Center, University Children's Hospital, Steinwiesstrasse 75, 8032 Zurich, Switzerland
d Children's Research Center, University Children's Hospital, Steinwiesstrasse 75, 8032 Zurich, Switzerland
e Department of Child and Adolescent Psychiatry, DRK Kliniken Berlin Westend, Spandauer Damm 130, 14050 Berlin, Germany

Objective

This study aims to evaluate the computerized **adaptive training program Calcularis** that offers children with developmental dyscalculia (DD) or math difficulties at an early stage the possibility to train especially those aspects of numerical cognition, in which they still need support. Calcularis disposes over a user model allowing flexible adaptation on the basis of the internally mapped learning and knowledge profile of the individual child.

Methods

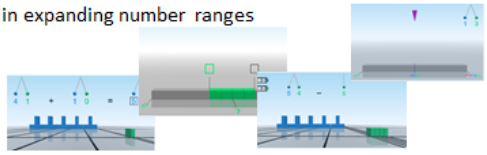
Study design and sample

- 67 children with DD or severe math difficulties aged 7;0 to 10;11 years (M = 8.85, SD = 0.85)
- 47 girls, 20 boys (χ^2 (1) = 10.88, p < .001)
- Inclusion criteria: IQ-scores within normal range (16th to 84th percentile) and low arithmetic-scores (<15th percentile)
- Children were randomly assigned to one of three groups (Calcularis group, waiting control group (WCG), control training group (spelling))
- Training duration: 6 weeks (5 sessions à 20 minutes per week)

	t_1	6-8 weeks	t_2
Calcularis group	t_1	Calcularis training	t_2
waiting group	t_1	waiting	t_2
control training group	t_1	control training	t_2

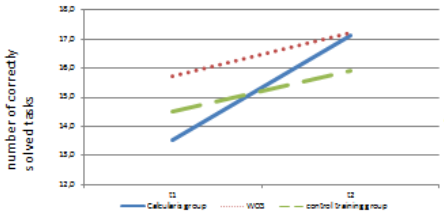
Structure of Calcularis

- Adaptive design
- Immediate response
- Training of **number representation**, formation and access to a **mental number line**, **arithmetic operations** in expanding number ranges

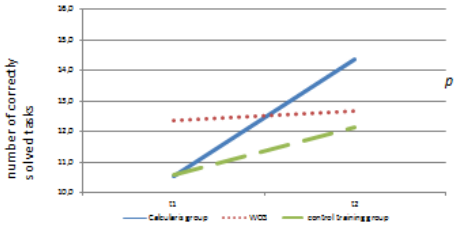


Results

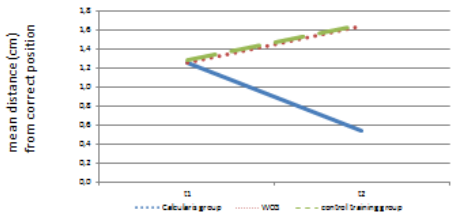
Addition (HRT)[1]



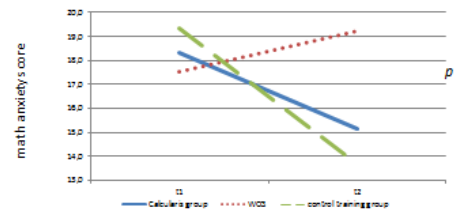
Subtraction (HRT)[1]



Position on mental number line (0-10)



Math anxiety (MAI)[2]



Conclusions

This study demonstrates that the adaptive training program Calcularis can be used effectively to support children with DD or math difficulties in their **numerical development** and to enhance **numerical cognition**. The results show that even after a rather short training period good effects with regard to **addition**, **subtraction** and **spatial number representation** were achieved. Furthermore, Calcularis led to a significant decrease of **math anxiety**, which may represent also an effect of increasing feelings of domain independent self efficacy, as the control training showed similar effects on math anxiety. Therefore we can strongly recommend implementing Calcularis into special need school programs and treatment procedures for children with developmental deficits in numerical cognition.

[1]Haffner, J., Baro, K., Parzer, P., & Resch, F. (2005). *Heidelberger Rechentest (HRT): Erfassung mathematischer Basiskompetenzen im Grundschulalter*. Göttingen: Hogrefe.
[2]Kohn, J., Richtmann, V., Rauscher, L., Kucian, K., Käser, T., Grond, U., Esser, G., & von Aster, M. (2013). Mathematikangstinterview (MAI) – erste psychometrische Gütekriterien. *Lernen und Lernstörungen*, 2(3), 177-189.

Contact: m.aster@drk-kliniken-berlin.de