Acta Psychologica 171 (2016) 118–127

Contents lists available at ScienceDirect

Acta Psychologica

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

Computer-based training for improving mental calculation in third- and fifth-graders

Sara Caviola [a,*], Giulia Gerotto [b], Irene C. Mammarella [b]

[a] Department of Psychology, University of Cambridge, Downing Street, Cambridge CB2 3EB, United Kindom
[b] Department of Developmental and Social Psychology, University of Padova, Via Venezia 8, 35131 Padova, Italy

ARTICLE INFO

Article history:
Received 4 June 2015
Received in revised form 6 October 2016
Accepted 13 October 2016
Available online xxxx

Keywords:
Mental calculation
Training
Mathematics

ABSTRACT

The literature on intervention programs to improve arithmetical abilities is fragmentary and few studies have examined training on the symbolic representation of numbers (i.e. Arabic digits). In the present research, three groups of 3rd- and 5th-grade schoolchildren were given training on mental additions: 76 were assigned to a computer-based strategic training (ST) group, 73 to a process-based training (PBT) group, and 71 to a passive control (PC) group. Before and after the training, the children were given a criterion task involving complex addition problems, a nearest transfer task on complex subtraction problems, two near transfer tasks on math fluency, and a far transfer task on numerical reasoning. Our results showed developmental differences: 3rd-graders benefited more from the ST, with transfer effects on subtraction problems and math fluency, while 5th-graders benefited more from the PBT, improving their response times in the criterion task. Developmental, clinical and educational implications of these findings are discussed.

1. Computer-based training for improving mental calculation in third- and fifth-graders

Learning different aspects of arithmetic is one of the main areas of academic achievement in which children often encounter difficulties, and the number of students with mathematical difficulties has greatly increased over the last 20 years (Swanson, 2011). Several studies (Geary, 2010; Lewis, Hitch, & Walker, 1994; Shalev & Gross-Tsur, 2001) indicate that 4–7% of the school-age population experience such difficulties in some form. Hence the growing interest in interventions to improve basic academic skills and reduce the number of children with mathematical difficulties. The literature on intervention programs to improve arithmetical abilities is still fragmentary, however, and - most importantly - the impact of previous interventions on math achievement is still not clear (Frank & Barner, 2012; Kucian et al., 2011). Previous studies, often based on the development of early arithmetic, tested interventions that can be classified by the type of task children were administered, in relation to the external magnitude of the representation of the numerical input (i.e. symbolic or non-symbolic representation; Butterworth, 2005; Dehaene, 2009), while they paid little attention to more complex numeracy skills, such as mental calculation (Obersteiner, Reiss, & Ufer, 2013). In this study, we developed and assessed two different types of training (strategy-based and process-based) in a controlled experimental setting with in 3rd and 5th grade schoolchildren. At these ages, children are starting to become familiar with both mental and written calculations of all four algorithms (in 3rd grade), and their skills are gradually consolidated and become more automatic and less demanding with school experience (by 5th grade).

1.1. Symbolic and non-symbolic number representations

During the early stages of numerical development, two non-verbal cognitive domains are responsible for the acquisition of the basic numerical processing skills: the exact and the approximate number systems (Butterworth, 2005; Halberda, Mazzocco, & Feigenson, 2008). The former precisely represents small numerosities, the latter approximately represents larger quantities. Both these cognitive systems rely initially on non-symbolic (i.e. non-verbal) codes, that are usually considered discrete for the exact representation, and discrete or continuous for the approximate system. With formal education, these codes become integrated with verbal (i.e. number words) and symbolic (i.e. Arabic digits) representations, and they provide the basis for subsequent numerical development (Feigenson, Dehaene, & Spelke, 2004). As regards arithmetical proficiency, a large body of evidence – based mainly on correlation and regression models (De Smedt, Noël, Gilmore, & Ansari, 2013; Sasanguie, Göbel, Moll, Smets, & Reynvoet, 2013) - confirms the relevance of the exact and approximate number systems in learning formal mathematics, with symbolic or non-symbolic representations of magnitude. Educational interventions have focused primarily on improving number system knowledge by means of magnitude comparison tasks, number-space mappings, number recognition, and

* Corresponding author.
E-mail addresses: sc2014@cam.ac.uk (S. Caviola), irene.mammarella@unipd.it (I.C. Mammarella).

http://dx.doi.org/10.1016/j.actpsy.2016.10.005

counting tasks. Most training programs have been implemented with kindergarteners or children from low-income backgrounds, and their effects have been apparent mainly on symbolic measures (Obersteiner et al., 2013; Toll & Van Luit, 2014). Very few studies have been conducted with older children or included a passive control group given no training, and none have attempted to enhance mental calculation skills directly by comparing process-based and strategy-based trainings.

1.2. Training approach: an overview

The existing literature distinguishes between strategy-based and process-based training schemes. Most existing cognitive training interventions aim to improve cognitive functions by teaching strategies (see for example, Carretti, Borella, & De Beni, 2007; Caviola, Mammarella, Cornoldi, & Lucangeli, 2009; Lustig & Flegal, 2008, for strategy-based training on working memory; or Verhaeghen, Marcoen, & Goossens, 1992, for a meta-analysis of strategy training programs for older adults). Strategic training typically involves identifying tasks in which a group of participants performs poorly and training them to use strategies that may help improve their performance. Another approach has been to train specific cognitive processes, without explicitly providing strategic training. These programs typically train participants on a set of tasks thought to load heavily on a specific cognitive process, and then measure transfer to a separate, untrained set of tasks also thought to load on the targeted process (see for example, Jaeggi, Buschkuehl, Jonides, & Perrig, 2008; Borella, Carretti, Riboldi, & De Beni, 2010, on working memory; Park, Gutchess, Meade, & Stine-Morrow, 2007, on older adults; Karbach, Strobach, & Schubert, 2015, for process-based training on mathematical achievement; and Schubert, Strobach, & Karbach, 2014, for a recent review on cognitive training interventions).

The same idea is often presented in the context of computer-assisted training programs (see e.g. Butterworth & Laurillard, 2010), but in the field of numerical processing it is also important to consider the distinction between symbolic (i.e. Arabic digits) and non-symbolic (i.e., non-verbal) representations. The types of training tested to date have consequently differed in several aspects, including the number format used (symbolic or non-symbolic code), and the instructions given to participants. Some researchers argue that the approximate number system (e.g. when students are asked to estimate the numerosities of large sets of objects and information is represented using discrete units such as dots) is particularly important to numerical development (Dehaene, 2009), and an adaptive game intervention (Wilson, Dehaene, Dubois, & Fayol, 2009) has been developed with a view to improving early numeracy. Other researchers claim instead that mathematical and arithmetical abilities rely on a “number module” that represents exact numerosities (Butterworth, 1999, 2005). Obersteiner et al. (2013) recently implemented one exact and one approximate version of the same computer-based training program (drawn from “The Number Race” by Wilson et al., 2006) and compared four groups of 6- to 7-year-old children: one group received only the exact training; one only the approximate training; one received both types of training in alternate sessions; and a control group received a language training. The results indicated that the groups receiving either one or the other of the two types of training improved in arithmetical performance by comparison with the other two groups.

Other research has shown that activities in preschool age (e.g., Clements & Sarama, 2007; Toll & Van Luit, 2014), kindergarten (e.g., Griffin, Case, & Siegler, 1994), or first grade (Fuchs et al., 2005) can substantially improve math performance. Although such tutoring activities are effective, not all students respond to them: the need for intensive remedial intervention persisted for a small percentage of children, even when preventive support services had proved generally effective (Fuchs et al., 2008).

Turning to symbolic representation, very few intervention studies have been conducted and the focus has been rather limited, addressing only basic facts or simple computation, and using drill and practice in brief intervention programs. In the area of mental calculation, Delazer et al. (2005) compared a strategic training with a training based on pure drill and repetition in adults. In the strategic training, the calculation problems varied in duration, and participants were asked what kind of strategy they used to solve each mental calculation, but no strategies for solving mental calculations were suggested. The results showed that accuracy improved more after the strategic training than after the drill and repetition training.

Taken together, these findings support the assumption that approximate or exact training can have a positive impact on tasks specifically related to those trained, but none of the studies clearly identified any transfer effects on arithmetical achievement, and the few studies that considered this aspect showed rather small effects of the training (Räsänen, Salminen, Wilson, Aunio, & Dehaene, 2009). That is why neither exact nor approximate training were used in the present study, which focuses instead on a strategic training for solving mental calculation problems, as compared with a process-based training, with a view to examining which type of training can enhance mental calculation skills and arithmetical achievement in children attending the 3rd and 5th grades of primary school.

A controversy in the field of education and teaching concerns how much instructional guidance needs to be provided in a learning environment (see Lee & Anderson, 2013 for a review). Learning conditions that introduce some degree of difficulty in the teaching provided appear to slow the learning rate, but to enable a better transfer than less difficult learning conditions (Bjork, 1994, Schmidt & Schmidt & Bjork, 1992). Several published studies have shown the superiority of direct instructions in mathematics (Carroll, 1994; Cooper & Sweller, 1987; Sweller & Cooper, 1985), while other research has suggested that students learn better in a discovery learning environment, in which they practice with their own strategies (Brunstein, Betts, & Anderson, 2009; Carpenter, Franke, Jacobs, Fennema, & Empson, 1998).

Within this scenario, it is important to note that the usefulness of a particular training could also be influenced by how students' levels of expertise/knowledge interact with the cognitive load of the tasks. Several researchers have said that the success of a particular training depends on the features of learners' cognitive processes, which depend on their personal domain-specific knowledge base (Blayney, Kalyuga, & Sweller, 2010; Kalyuga, Law, & Lee, 2013). A training that reduces the cognitive load of new knowledge for students (e.g. by providing plenty of instructions or by breaking down complex task guidelines into a number of intermediate steps) might be less effective for more skilled students (the expertise reversal effect), whereas such expert students may learn better without guidance. In other words, the additional instructions that are valuable to less expert students could impair the learning of the more expert (Lee & Kalyuga, 2014).

1.3. Research questions and hypothesis

The aims of the present study were: to develop two types of training on mental additions, one based on teaching strategy use and the other on repeated practice (i.e., process-based), both within a carefully controlled setting; and to assess any effects on the criterion task (i.e. mental addition problems), and any transfer effects, not only on tasks closely linked to the arithmetical domain (such as mental subtraction), but also on others near (math fluency) and far (numerical reasoning).

We focused on two main research questions:

1) Do strategy and process-based training interventions have different effects on the mental calculation skills of children attending primary school?
2) Are there specific differences in relation to the child's age and the type of training administered?

For both types of training, learning gains were expected to be greater than in an untrained control condition. Different effects were also expected as a result of the different nature of the two types of training.

The study was conducted on 3rd- and 5th-grade schoolchildren, primarily because few other studies have considered these age groups from a developmental perspective, and also because specific benefits could emerge in relation not only to the type of training proposed, but also to the children's educational level and cognitive competences. Children at the ages considered here might be representative of two different, sensitive points in the development of mental calculation skills that can contribute to determining different patterns in the training effects: children in 3rd grade have yet to completely master mental calculation, but by 5th grade they will have started to automatize and develop efficient mental calculation skills (Baroody & Dowker, 2003). Thus, consistently with the expertise reversal effect (Kalyuga, 2007; Kalyuga, Ayres, Chandler, & Sweller, 2003), we would expect the inexperienced learners (3rd-graders) to benefit from the strategic training more than the experienced learners (5th-graders).

2. Method

2.1. Participants

Primary school children from 6 public elementary schools distributed evenly between the city and the countryside that serve families from a broad range of socioeconomic backgrounds were invited to participate. In all, 225 children were tested. Our sample was recruited from two different cohorts, based on year of schooling, i.e., third-graders ($N = 120$, mean age = 101.75 months, SD = 4.73), and fifth-graders ($N = 105$, mean age = 126.14 months, SD = 4.49). Six children did not complete all the sessions and their data were not considered in this report. The final sample thus included 219 children, who were randomly grouped as follows: the strategic training (ST) group included 41 third- (15 M, 26 F), and 35 fifth-graders (10 M, 25 F); the process-based training (PBT) group contained 38 third- (21 M, 17 F) and 34 fifth-graders (14 M, 20 F); and 38 children in third (19 M, 19 F) and 33 in fifth grade (15 M, 18 F) were assigned to a passive control (PC) group.

The participants' parents/caregivers were contacted via the school administrator to explain the purpose of the study and the procedures involved, and to ask them to sign a consent form on behalf of their children if they agreed to their inclusion in the study. The procedures adopted to obtain the informed written consent of the parents/caregivers and the children's verbal assent were consistent with the APA guidelines. An eligibility criterion was that the children were not being considered for, or already the object of an individualized education plan for children with demonstrated special needs at the time of our study.

The children in the three groups were matched for gender and for measures of verbal and spatial reasoning, based on two subtests of the Primary Mental Ability Test (PMA) (Thurstone & Thurstone, 1963). In particular, verbal reasoning skills were assessed by using the Verbal Meaning subtest: children had to identify one among five possible words that shared the meaning of a target word. Spatial reasoning abilities were assessed with the Spatial Relation subtest: the task involved identifying which "irregular" shape could be rotated and merged with the target to form a perfect square.

Specifically, there were no gender-related differences for either the third-graders $\chi^2 = 2.96$ $p = .23$, or the fifth-graders $\chi^2 = 2.24$ $p = .33$. There were also no differences in the scores obtained by the third-graders $F(2, 114) = 1.15$, $p = .32$, $\eta^2 = .02$, and fifth-graders $F(2, 99) < 1$, $\eta^2 = 0.007$, in the Verbal Meaning subtest, or in the Spatial Relations subtest, $F(2, 114) < 1$, $\eta^2 = 0.006$ and $F(2, 99) = 1.88$, $p = .16$, $\eta^2 = .04$, respectively. The characteristics of the groups are summarized in Table 1.

Table 1
Characteristics of the three groups: strategic training (ST), process-based training (PBT) and passive controls (PC).

	Grades	Age (months)		Verbal meaning (PMA)		Spatial relations (PMA)	
		M	SD	M	SD	M	SD
ST	3rd ($N = 41$)	100.95	4.71	11.58	3.11	5.68	3.11
	5th ($N = 35$)	124.62	4.99	10.61	3.31	5.26	3.31
PBT	3rd ($N = 38$)	101.87	4.46	11.34	2.37	5.61	2.37
	5th ($N = 34$)	125.43	4.21	13.91	1.84	6.60	2.07
PC	3rd ($N = 38$)	102.50	3.64	13.71	2.43	6.62	2.05
	5th ($N = 33$)	124.84	4.47	13.51	1.54	5.69	2.52

2.2. Materials

2.2.1. Criterion training task

2.2.1.1. Complex mental addition problems. 16 multi-digit addition problems (with three- or two-digit operands), half with and half without carrying over in the units. The material was prepared using E-Prime software (Psychology Software Tools, Inc., Pittsburgh, PA, USA) on a 15-inch computer screen. The children were seated in front of the screen, placed about 60 cm away, and were asked to solve addition problems, all presented horizontally. There were no time limits and children only had one chance to indicate the solution for each problem. Accuracy and response time were considered as dependent variables. Cronbach's alpha = .71.

2.2.2. Near and far transfer training tasks

2.2.2.1. Complex mental subtraction problems. The children were asked to solve 16 multi-digit subtraction problems (with three- or two-digit operands), half with and half without borrowing in the units. All the problems were presented horizontally. There were no time limits and children only had one chance to indicate the solution for each problem. Accuracy and response times were considered as dependent variables. Cronbach's alpha = .73.

2.2.2.2. Math fluency in additions and subtractions (Caviola, Gerotto, Lucangeli, & Mammarella, in press). The measure of math fluency in additions involved one paper-printed page with 24 multi-digit additions and one with 24 subtractions. The children were given 2 min to complete each page and were asked to solve the problems as quickly and accurately as possible. The number of additions and subtractions solved correctly was recorded. Cronbach's alpha was >.83 for both subtests.

2.2.2.3. Numerical reasoning: (adapted from Thurstone & Thurstone, 1963). This task assesses the ability to discover rules and apply them to numerical reasoning activities. It is a written test in which the children had to choose which number was missing from a set from four options, e.g. 40, 45, 50, 55, ___, 65 (options: 40, 60, 70, 45). It included 10 sets of items and the children were allowed 5 min to complete the task. The score was the sum of the correct answers. Cronbach's alpha = .74.

2.3. Procedure

All participants attended pre- and post-training sessions on two separate occasions: one as a group in their classrooms, which lasted about 40 min (when they performed the math fluency in additions and subtractions, and the numerical reasoning tasks); the other lasting about 20 min and attended individually, when the children were administered the other tasks (mental addition problems, mental subtraction problems). For each of the tasks presented, two parallel versions were devised and administered in a counterbalanced order before and after the training sessions.

The training programs followed a double-blind controlled design. The experimenters were three female Italian postgraduate students trained by the authors. Two experimenters carried out the pre- and post-training assessments, while the third managed both of the training conditions. The experimenters completing the assessments did not know which group the children were in, and the experimenter conducting the training programs was unaware of the study hypotheses.

2.3.1. Mental calculation training

Both the strategic and the process-based training interventions consisted of three sessions of about 1 h in which participants were given 32 addition problems for each session. The same addition problems were presented for both types of training. The difficulty of the addition problems varied across sessions (see Table 2) and in all the sessions half of them involved carrying over.

Each training session started with two examples to make the children more confident with the task and the computerized program.

In the strategic training (ST) group, children were shown 26 of the 32 addition problems with strategy cues: children could follow the strategic instructions on the computer screen, which explained how to break down the calculation into 2 or 3 easier steps. A strategy was suggested for each addition problem. We used examples of transformation strategies, as defined by LeFevre, DeStefano, Penner-Wilger, and Daley (2006), which include the following: breaking down one or both operands; rounding up/down; and fact retrieval for other operations (i.e. $85 + 13 = 85 + 10 = 95; 95 + 3 = 98$). The children had to compute the partial results correctly in order to arrive at the solution. The last 6 problems were presented according to the procedure used in the process-based training and children were asked to use the previously-suggested strategies. There were no time constraints.

The process-based training (PBT) training was introduced as an exercise on mental addition problems and children were asked to solve the operations (which were exactly the same addition problems as for the ST group) as quickly and accurately as possible. The children were given no instructions regarding the use of strategies. There were no time constraints for the task's completion and the children were asked to do their best. Examples of both training procedures and types of problem are given in Table 2. Each session lasted about an hour, with slight differences depending on each child's competence.

Children could have 3 attempts at solving each problem, in both the ST and the PBT conditions. They received feedback for each answer they gave: if their answer was correct, they were invited to press the right-arrow key to go on to the next problem; if not, they were asked to press the left-arrow key and to try and solve the same problem again.

The three training sessions were scheduled with an interval of 3 days between them, and the post-training sessions were completed within 5 days after the last training session.

Participants in the passive control (PC) group only completed the pre- and post-test assessments.

2.3.2. Fidelity implementation

During the training sessions in the experimental condition, the trainer kept a daily journal of the activities involved in each session. The computer recorded the percentage of correct answers and the median response times for each session. In the ST group, 3rd graders correctly solved 91%, 80% and 78% of the problems in the three successive sessions (the median response time ranged from 11 to 12 s); whereas 5th-graders solved 88%, 84% and 76% of addition problems, respectively, within a temporal window of 8 to 10 s. Children in 3rd grade assigned to the PBT group completed 87%, 76% and 84% of the additions correctly (the median response time ranged from 13 to 18 s); and 5th-graders in this group correctly solved 90%, 82% and 86% of the addition problems in the three sessions with a median response time ranging from 8 to 10 s.

2.4. Data analyses

First, preliminary analyses were conducted at the pre-training session (distinguishing the 3rd- from the 5th- graders) to ensure that participants in the ST, PBT and PC groups had a comparable performance at the baseline.

Second, mixed-design ANOVAs, 3 (Group: ST, PBT, and PC) × 2 (Grade: third, fifth) × 2 (session: pre-test, post-test) were run separately on all measures of interest to examine the training-related gains in more detail. Post hoc analyses were corrected for multiple comparisons

Table 2
Session content for the strategic training (ST) and process-based training (PBT) groups.

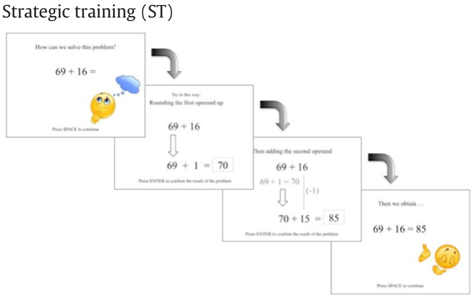
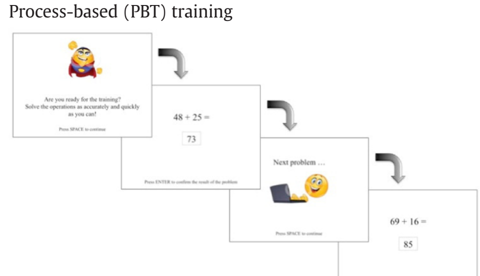
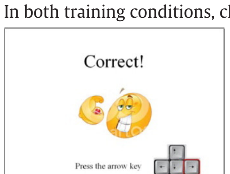
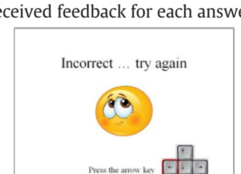
Details of the training: procedure, instruction, feedback and materials		
Examples of training activities	Strategic training (ST) 	Process-based (PBT) training 
Instructions	Children solved 26 of the 32 addition problems with instructions on strategy use. The last 6 problems were presented as in the PBT training.	In the PBT training children were asked to solve the operations as quickly and accurately as possible.
Problems presented	1st session. 32 addition problems: half of them comprising 2 digits +1 digit, the other half 2 + 2 digits; half of the problems required no carrying over, and the other half involved a carrying over procedure on the units. 2nd session. 32 addition problems: half of them comprising 2 + 2 digits, the other half 3 + 2 digits; half of the problems required no carrying over, and the other half involved a carrying over procedure on the units. 3rd session. 32 addition problems: 8 comprising 2 digits +1 digit, 16 involved 2 + 2 digits, and 8 involved 3 + 2 digits; half of the problems required no carrying over, and the other half involved a carrying over procedure on the units.	
Examples of feedback	In both training conditions, children had up to 3 attempts to solve each problem. They received feedback for each answer they gave. 	

Table 3
Descriptive statistics: means (M) and standard deviations (SD) for measures of interest by group (strategic training [ST], process-based training [PBT], and passive control [PC]), and by grade (3rd and 5th).

		3rd-graders			5th-graders		
		ST M (SD)	PBT M (SD)	PC M (SD)	ST M (SD)	PBT M (SD)	PC M (SD)
Accuracy - additions	Pre-test	11.95 (2.97)	11.94 (3.09)	11.92 (2.83)	12.66 (2.44)	12.41 (2.57)	13.09 (2.30)
	Post-test	13.59 (2.08)	13.24 (1.94)	13.09 (2.30)	13.80 (1.80)	12.88 (2.41)	12.43 (2.61)
Response times - additions	Pre-test	15,739 (6859)	15,286 (5175)	15,669 (5834)	10,600 (3045)	10,775 (3806)	11,491 (3155)
	Post-test	15,039 (8163)	17,086 (7462)	14,607 (5510)	8783 (3488)	8456 (3546)	11,115 (4255)
Accuracy - subtractions	Pre-test	11.29 (2.47)	10.55 (3.67)	11.26 (3.33)	12.17 (2.70)	12.18 (2.72)	12.06 (3.17)
	Post-test	12.41 (2.47)	11.63 (2.39)	10.05 (3.24)	11.49 (2.78)	11.32 (3.10)	11.21 (2.51)
Response times - subtractions	Pre-test	18,338 (8351)	18,074 (7002)	19,280 (6849)	12,368 (2849)	12,191 (3831)	14,190 (4917)
	Post-test	21,153 (8000)	27,238 (14342)	20,571 (7483)	11,972 (3679)	11,546 (4568)	15,561 (6347)
Math fluency - additions	Pre-test	10.59 (2.05)	9.34 (2.80)	9.63 (3.16)	16.77 (4.85)	17.29 (3.35)	16.06 (2.93)
	Post-test	15.17 (4.53)	12.45 (3.19)	10.95 (2.50)	18.00 (3.89)	18.71 (3.97)	17.15 (4.18)
Math fluency - subtractions	Pre-test	8.63 (2.55)	8.05 (2.57)	8.08 (3.35)	12.69 (3.91)	13.85 (3.29)	12.15 (4.36)
	Post-test	10.34 (4.06)	9.53 (3.12)	8.89 (2.39)	14.26 (4.97)	15.18 (4.87)	13.18 (4.10)
Numerical reasoning	Pre-test	7.07 (2.26)	6.71 (1.87)	7.26 (1.67)	8.43 (1.01)	8.15 (1.41)	8.12 (1.83)
	Post-test	7.51 (1.86)	7.05 (1.47)	7.42 (1.64)	8.60 (1.03)	8.44 (1.19)	8.27 (1.77)

Accuracy - additions and subtractions: number of correct responses; math fluency - additions and subtractions: number of correct responses; numerical reasoning: number of correct responses. Response times for additions and subtractions are expressed in milliseconds.

using Bonferroni's adjustment and were not performed when the interaction was not significant.

Finally, to better examine differences among the groups, Cohen's d values were computed to clarify the relative differences in the effects between the experimental groups (ST and PBT) and the control group (PC). In other words, the post-test differences between the groups were divided by the pooled standard deviation for both age groups. Then all d values were corrected for small sample bias using the Hedges and Olkin (1985) correction factor (d').

3. Results

3.1. Preliminary analysis

Table 3 summarizes the descriptive statistics, by school grade and training group. The preliminary analyses (one-way ANOVAs) revealed no significant differences between the groups at the pre-training assessment (3rd-graders: Complex mental addition problems – accuracy $F(2, 114) < 1$; complex mental addition problems – response times $F(2, 114) < 1$; complex mental subtraction problems – accuracy $F(2, 114) < 1$; complex mental subtraction problems – response times $F(2, 114) < 1$; math fluency - additions $F(2, 114) = 2.32$, $p = 0.11$ $\eta^2 = 0.04$; math fluency - subtractions $F(2, 114) < 1$; numerical reasoning $F(2, 114) < 1$; 5th-graders: complex mental addition problems – accuracy $F(2, 99) < 1$; complex mental addition problems – response times $F(2, 99) < 1$; complex mental subtraction problems – accuracy $F(2, 99) < 1$; complex mental subtraction problems – response times $F(2, 99) = 2.66$, $p = .09$ $\eta^2 = .05$; math fluency additions $F(2, 99) < 1$; math fluency subtractions $F(2, 99) = 1.70$, $p = .19$ $\eta^2 = .03$; numerical reasoning $F(2, 99) < 1$).

3.2. Training-related gains

We ran mixed-design, 3 (Group: ST, PBT, and PC) × 2 (Grade: third, fifth) × 2 (session: pre-test, post-test) ANOVAs separately on all measures of interest: accuracy measures included the sum of the correct answers in each task, while for response times only correct answers were considered. All the effects of these analyses are shown in Tables 4 and 5, and only significant effects are reported in the text.[1]

[1] For the sake of simplicity, the two-way interactions are discussed only when the three-way interactions were not significant.

3.3. Criterion training task

3.3.1. Complex mental addition problems

Regarding accuracy, the main effect of grade was significant, showing that children in third grade performed significantly worse than those in fifth grade (12.5 vs. 13.0, $p = .043$). The main effect of session was also significant, showing that accuracy improved in the post-training assessment (12.3 vs. 13.2, $p < .0001$). However, the interaction between group and session, intended to highlight improvements from pre- to post-training by group, was not significant (all $p_s > .33$, see Table 4).

Concerning response times, the main effect of grade was significant, meaning that 5th-graders were faster than 3rd-graders (10.2 s vs. 15.5 s, $p < 0.0001$); the main effect of session was also significant: at post-test children solved the problems more quickly (13.2 s vs. 12.5 s, $p = .044$). The session × grade interaction was also significant ($p = .041$, see Table 4). Statistically larger improvements from pre- to post-training clearly emerged only in the three-way interaction session × group × grade ($p = 0.030$), underscoring the additional role of age differences in modulating the effects of training. Post hoc analyses only showed a difference in terms of response times for 3rd-graders in the PBT group, revealing that younger children were significantly slower after the training than at the pre-training assessment (15 s vs. 17 s, $p = .043$), while response times did not change in the ST ($p = .41$) and PC ($p = .23$) groups. Among the 5th-graders, on the other hand, improvements in performance were larger in both the training groups: both the ST (10.5 s vs. 8.7 s, $p = .049$) and the PBT (10.7 s vs. 8.4 s, $p = .014$) groups worked faster after the training than at the pre-training assessment showing a clear reduction in terms of rapidity, while the PC group's response times ($p = .69$) did not change. Finally, for both training groups, and at both pre- and post-test sessions, 5th-graders were generally faster than 3rd-graders ($p_s < .013$), as expected (given the age difference). Non-significant effects are reported in Table 4.

3.4. Near and far transfer training tasks

3.4.1. Complex mental subtraction problems

Regarding accuracy, there were no significant main effects, but - more importantly - differences emerged in terms of the interactions, which were significant for session x grade ($p = .008$, see Table 4) and session × group ($p = .030$), indicating that training gains differed by group. The three-way interaction session × group × grade was also significant ($p = .047$), showing that the efficacy of the training was differently affected by both group and age: post hoc analyses indicated that

Table 4
Results of the mixed-design 2 (grade: third and fifth grade) × 3 (group: ST, PBT and PC) × 2 (session: pre- and post-test) ANOVAs for the computerized tasks: the criterion task (complex mental addition problems) and the nearest transfer task (complex mental addition problems).

	Mental additions accuracy			Mental additions RTs			Mental subtractions accuracy			Mental subtractions RTs		
	F	p	η_p^2	F	p	η_p^2	F	p	η_p^2	F	p	η_p^2
Grade (G)	**4.142**	**.043**	**.019**	**71.484**	<**.0001**	**.251**	2.607	.108	.012	**91.972**	<**.0001**	**.302**
Group (Gr)	.610	.544	.006	.386	.680	.004	1.488	.228	.014	1.300	.275	.012
Session (S)	**20.406**	<**.0001**	**.087**	**4.089**	**.044**	**.019**	1.229	.269	.006	**16.570**	<**.0001**	**.072**
G * Gr	1.174	.311	.011	1.583	.208	.015	.802	.450	.007	**4.085**	**.018**	**.037**
G * S	1.399	.238	.007	**4.230**	**.041**	**.019**	**7.186**	**.008**	**.033**	**15.001**	<**.0001**	**.066**
Gr * S	1.477	.231	.014	.623	.538	.006	**3.580**	**.030**	**.033**	**3.194**	**.043**	**.029**
G * Gr * S	.198	.820	.002	**3.552**	**.030**	**.032**	**3.099**	**.047**	**.028**	**6.707**	**.001**	**.059**

Note: G = grade; Gr = group; S = session.

the third-graders' accuracy improved from the pre- to the post-training sessions in the ST (11.2 vs. 12.4, $p = 0.021$) and PBT (10.5 vs. 11.6, $p = .033$) groups, while it deteriorated in the PC group (11.2 vs. 10, $p = .017$). Considering only the post-training sessions for the children in third grade, a significantly greater increase in the number of correct answer was seen in the ST and PBT groups compared with the PC group ($p = .001$ and $p = .04$, respectively). There were no significant differences between the groups for the children in 5th grade, indicating that neither experimental training intervention enhanced the older children' performance from the pre-training to the post-training session. As a consequence, no differences emerged between the groups for the 5th-graders in terms of post-training performance.

Concerning response times, the main effect of grade was significant, meaning that 5th-graders performed faster than 3rd-graders (13 vs. 21 s, $p = .0001$). The main effect of session was significant too, showing an increase in the response times at the post-test session (16 vs. 18 s, $p = .001$). All the two-way interactions were significant: session × grade ($p < .0001$), group × grade ($p = .018$), and session × group ($p = .043$). The three-way interaction session × grade × group was also significant ($p = .001$) and post hoc analyses revealed that 3rd-graders were slower after the training (compared to the pre-training session) in both the ST ($p = .029$) and the PBT groups ($p = .0001$), whereas the PC ($p = .33$) group's response times remained the same. When only the post-training session was considered, however, the 3rd-graders attending the PBT groups had significantly slower response times than the children assigned to the ST group ($p = .004$) or PC group ($p = .002$), suggesting that the PBT training had a stronger effect on response times than the ST training. No significant differences emerged for the 5th-graders, indicating that response times for complex mental calculation problems did not change after the training in 5th-graders, and no differences consequently emerged between the groups at the post-training sessions. All the others non-significant effects are reported in Table 4.

3.4.2. Math fluency - additions and subtractions

For math fluency in additions the main effect of grade was significant, with 5th-graders performing better than 3rd-graders (11 vs. 17, $p = .0001$, see Table 5). The main effect of group was significant ($p = .006$), the ST group performing better than the PC group ($p = .004$), but no better than the PBT group ($p = .56$). The main effect of session was also significant (13.2 vs. 15,4, $p = .0001$), showing an increase in the number of problems solved at the post-test training session. The session × group ($p = .006$), session × grade ($p = .001$), and grade × group ($p = .038$) interactions were significant too. Importantly, the three-way interaction session × grade × group was significant as well ($p = .013$), showing a significant pretest-posttest increase in the number of correct answers, and modulated by both group and age. First, considering the changes from the pre- to the post-training sessions, our results showed that performance improved in 3rd-graders in the ST ($p = .0001$), PBT ($p = .0001$), and PC ($p = .01$) groups, and the same was true of the groups of 5th-graders: ST ($p = .023$), PBT ($p = .01$) and PC ($p = .05$). At the post-test, 3rd-graders in the ST groups also obtained significantly more correct answers than those assigned to the PBT group ($p = .005$) or the PC group ($p < .001$). No significant differences emerged for the 5th-graders (see Table 5).

Our results for math fluency in subtractions revealed only a main effect of grade, meaning that 5th-graders performed better than 3rd-graders (13.5 vs. 9, $p < .0001$), and a main effect of session, showing that accuracy was better at the post-training assessment (10.5 vs. 12, $p < .0001$). No other effects were statistically significance (all $p_s > .33$).

3.4.3. Numerical reasoning

The main effect of grade was significant, showing that 5th-graders outperformed 3rd-graders (7.1 vs. 8.3, $p < .0001$); as in math fluency for subtractions, there was also a main effect of session, showing a slight increment in accuracy at the post-training assessment (7.6 vs. 7.8, $p = .024$). There were no other significant effects (all $p_s > .40$, see Table 5).

3.5. Training-related benefits based on Cohen's d in terms of differences among groups

Finally, to clarify the group differences, and particularly to highlight the most useful experimental training in each age group, effect sizes were calculated for all tasks by comparing post-test measures across groups. Fig. 1 shows Cohen's d values (expressing the effect size of the

Table 5
Results of the mixed-design 2 (grade: third and fifth grade) × 3 (group: ST, PBT and PC) × 2 (session: pre- and post-test) ANOVAs for the near transfer task (math fluency additions and subtractions) and far transfer task (numerical reasoning).

	Math fluency additions			Math fluency subtractions			Numerical reasoning		
	F	p	η_p^2	F	p	η_p^2	F	p	η_p^2
Grade (G)	**196.429**	<**.0001**	**.480**	**104.548**	<**.0001**	**.329**	**37.078**	<**.0001**	**.148**
Group (Gr)	**5.270**	**.006**	**.047**	2.132	.121	.020	.928	.397	.009
Session (S)	**97.196**	<**.0001**	**.313**	**39.015**	<**.0001**	**.155**	**5.183**	**.024**	**.024**
G * Gr	**3.313**	**.038**	**.030**	1.478	.230	.014	.720	.488	.007
G * S	**16.670**	<**.0001**	**.073**	.003	.955	<.0001	.222	.638	.001
Gr * S	**5.319**	**.006**	**.048**	.990	.373	.009	.209	.812	.002
G * Gr * S	**4.446**	**.013**	**.040**	.078	.925	.001	.129	.879	.001

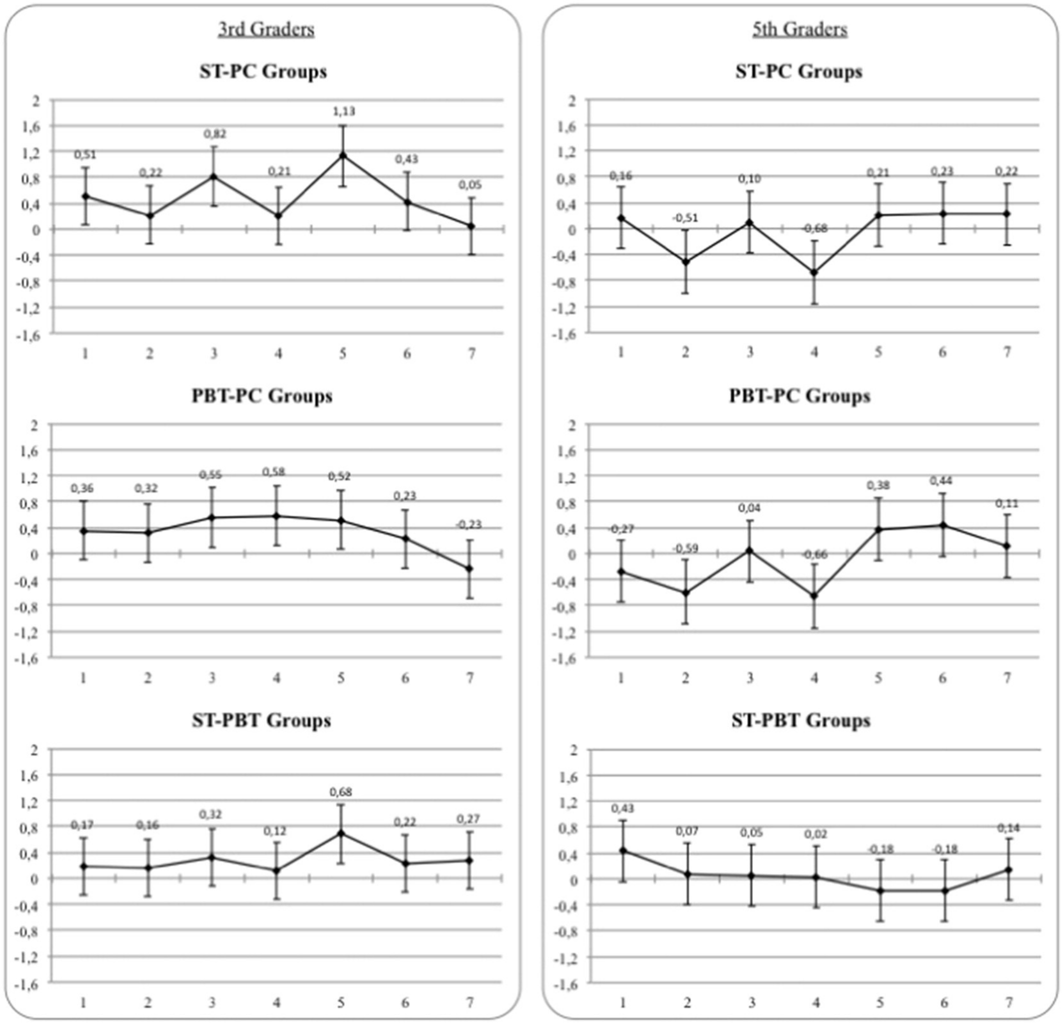


Fig. 1. Comparison between Cohen's d values for third-graders (left panel) and fifth-graders (right panel) in the post-test measures by group: strategic training (ST), process-based training (PBT) and passive control (PC). Note. The figure only shows Cohen's *d* values because the d_{unb} was largely redundant. 1 = accuracy - additions; 2 = response times - additions; 3 = accuracy - subtractions; 4 = response times - subtractions; 5 = math fluency – additions; 6 = math fluency – subtractions; 7 = numerical reasoning.

comparisons) for the specific and transfer effects detectable between the third- and fifth-graders in the experimental and control groups. Effect size was also measured with an unbiased variant of Cohen's *d* that accounts for small sample sizes (d_{unb}) (using the Hedges and Olkin, 1985 correction factor; see also Grissom & Kim, 2012): unbiased d_s values have only been reported when they differed from the regular d_s values. According to Cohen (1988), values in the range of $d = 0$; .5 would indicate a small effect size; values between $d = 0.5$; .8 a medium effect size; and values of $d > .8$ a large effect size.

The comparison between the ST and PC groups clearly highlighted different effects in the two age groups considered: the ST training showed medium effects in the accuracy measures for addition and subtraction problems (Cohen's $d = .50$ and .81 respectively) in 3rd-graders; similar effect sizes were seen in 5th-graders for the response times in addition and subtraction problems (Cohen's $d = -.51$ and $-.68$ respectively).

Concerning the benefit deriving from the PBT training (comparing the PBT and PC groups), the patterns showed different effects in the two age groups: in third grade, children reached only medium effects in the near transfer measures, i.e. accuracy in the subtraction problems (Cohen's $d = .55$) and in math fluency for additions (Cohen's $d = .52$); for the older children, on the other hand, the benefit was apparent in the response times for both additions and subtractions (Cohen's $d = -.55$ and $-.66$ respectively).

Finally, the comparison between the ST and PBT training showed only a medium effect size for the near transfer measure of math fluency for additions in 3rd-graders (Cohen's $d = .68$). No differences emerged in terms of effect sizes in 5th grade.

4. Discussion

The main aim of the present study was to analyze the effect of two different types of training (i.e., strategy-based and process-based) on the mental calculation skills of children attending the 3rd and 5th grades of primary school. The training involved three sessions in which children were presented with addition problems of increasing complexity. At pre- and post-training assessments, the children were presented with a criterion task involving complex addition problems, a nearest transfer task with complex subtraction problems, two near transfer tasks testing their math fluency (additions and subtractions), and a far transfer task relating to numerical reasoning.

Based on our hypotheses, we expected gains from both kinds of training by comparison with the passive control group, together with age-related differences. Looking at the differences between pre- and post-training in the criterion task, our data show an influence of the training in both the experimental groups, but only for speed of execution. Specifically, while 3rd-graders in the PBT group became slower, 5th-graders in both the ST and the PBT groups became faster than children in the PC group. In addition, comparing group differences at post-training revealed that the complex mental addition problems administered to the ST group showed a medium effect size in the 3rd-graders' accuracy, i.e. at the post-training assessment they became more accurate in solving mental addition problems. In 5th-graders, medium effect sizes were observed for the response times, indicating that children in both the ST and the PBT groups became faster than controls (Cohen's $d = .51$ and .55, respectively).

As for the near and far transfer effects, the third-graders assigned to the ST and PBT groups improved in terms of accuracy when performing complex subtractions, although the ST group showed a larger effect size (Cohen's $d = .81$) at post-training compared with the PC group. In terms of response times, 3rd-graders assigned to the PBT training became slower after the training than children assigned to the ST group, showing a weaker benefit of the process-based training. In addition, 3rd-graders showed a significant effect on math fluency in additions (with a large effect size for 3rd-grade children in the ST group). Finally, there were no far transfer effects on numerical reasoning. In short, both types of training were effective in the 3rd-graders, the ST slightly more so than the PBT.

A different pattern of results emerged for the 5th-graders, showing specific effects in the ST and PBT groups in terms of response times for complex subtractions, with children completing the calculations significantly faster. No specific effects were observed, however, on the math fluency tasks, or on numerical reasoning. Thus, the only specific effect of the two types of training on 5th-graders concerned response times for complex mental calculations, and this effect was slightly greater for the children in the PBT group.

To sum up, our findings seems to differ in the children assigned to the ST and PBT groups, depending on their age. Developmental differences emerged in relation to the efficacy of the two types of training. Children in third grade showed more evident transfer effects than those in fifth grade (for near tasks, at least). It is worth noting that children in 3rd grade have yet to completely master mental calculation, but by fifth grade they have started to automatize and develop efficient mental calculation skills (Baroody & Dowker, 2003). Also, given the expertise reversal effect (Kalyuga, 2007; Kalyuga et al., 2003), we expected a greater benefit of the strategic training in inexperienced learners (3rd-graders) than in experienced learners (5th-graders), and - judging from our results - 3rd-graders benefited more from the ST than from the PBT, while it seems that the opposite was true of the 5th-graders. This situation has been seen in many domains, including multimedia learning (Mayer & Mayer & Sims, 1994), probability calculation (Renkl, 1997), and logical programming (Kalyuga, Chandler, Tuovinen, & Sweller, 2001). According to Lee and Anderson (2013), the expertise reversal effect is enhanced when environment learning and cognitive ability are well matched. In particular, children acquire more robust declarative representations if rules are explained than if they are just asked to apply them. In our PBT training, it may be that experienced learners (5th-graders) were able to formulate strategies in the PBT condition because they already possessed the necessary cognitive skills. Our findings concerning response times and accuracy would support these assumptions, i.e. the 3rd-graders became more accurate, the 5th-graders quicker to respond. In addition, Titz and Karbach (2014) recently suggested, albeit in different types of training, that strategic training produced magnification effects (thereby augmenting individual differences), in the memory domain at least, whereas process-based training (focusing on working memory and executive functions, for example) promotes compensation effects that reduce individual differences. As a consequence, participants with an initially worse performance (i.e., 3rd -graders in our study) would benefit more from strategic training interventions. The differences seen in the transfer effects can also be explained on the grounds of age-related differences. In Italy, children in third grade start learning complex written additions and subtractions, and simple multiplications, while decimal numbers and fractions are taught in fifth grade. Our training with additions thus had an effect on both additions and subtractions in the younger children, but no clear transfer effects in the older children (who became faster in performing both additions and subtractions, but did not benefit in terms of accuracy from the three training sessions, possibly because they already knew the addition and subtraction algorithm well). Finally, we found no far transfer effects on a task measuring numerical reasoning: this may be because the children were administered only a very short, specific training on mental addition problems. In other words, strategic training seems to promote flexibility and the applicability of the strategies conceived to other contexts (i.e. math fluency tasks, paper and pencil time-based tasks) in 3rd-graders at least, whereas process-based training reduced the 5th-graders' response times in the criterion task, and helped them to develop more automatic mental calculation skills, without any transfer to other tasks.

The originality of our study lies in that: the children were trained on complex addition problems within a carefully controlled setting; and two types of training were compared, one in which specific strategies were suggested, and the other involving repeated practice (i.e., process-based). In the literature, most training activities were proposed for preschoolers (Clements & Sarama, 2007; Toll & Van Luit, 2014), in kindergarten (Griffin et al., 1994), or in first grade (Fuchs et al., 2005), and most of the research focused on exact or approximate non-symbolic representation (Obersteiner et al., 2013). Few intervention studies have been conducted on symbolic representation, addressing only basic facts or simple calculations (Delazer et al., 2005; Frank & Barner, 2012). Recently, Räsänen (2015) systematically described the historic development of computer-assisted interventions since the 1970s, and the Author analytically reported the results of different reviews on the efficacy of such interventions. Although some studies had negative or null effects, positive effects were generally detected, albeit with a small effect size (around .3 in Cohen's d). This leads to the conclusion that effect sizes should be considered in the research context as a whole, such as when the results may have important "practical" implications, or when an ongoing situation is investigated (i.e. the mental calculation learning that should precede the procedural learning of complex written arithmetic) (Prentice & Miller, 1992a,b,c; Stukas & Cumming, 2014). Our findings should thus be interpreted within this framework.

Some limitations of our study must be acknowledged, however. The first concerns the duration of our training (only three sessions lasting 1 h each). This was enough to demonstrate some benefits of the two types of training, but a longer period of training would probably generate stronger results, particularly as regards any transfer effects. Other limitations concern the fact that the children were not presented with non-symbolic tasks at either the pre- or the post-training assessments, and no other measures of their calculation skills were considered (such as written calculation). To better elucidate the benefit of any training, further research should therefore test non-symbolic representations of numbers, and mathematical achievement too. Another weak point of our study lies in the absence of a follow-up session, which limits the generalizability of our results in terms of the duration of any effects of the training. Unfortunately, the schools and the two cohorts of children were unavailable for any such follow-up.

The present study has some clinical implications: intervention programs should be designed bearing the baseline level of the children's knowledge in mind in order to improve their basic academic skills and reduce the number of children with mathematical difficulties. The educational implications that can be drawn from our findings could help teachers and educators to design the most useful type of training for different age groups of children. In particular, our findings reinforce the conviction that inexperienced learners benefit more from clear instructions concerning the best strategy to use in solving mental calculations, while experienced learners benefit more from repeated practice that enables them to develop their own strategies and reduce their response times. The difficulties encountered by students in arithmetic pose a challenge for teachers and educators, and a better understanding of the cognitive processes involved in mental calculations, and of the efficacy of different types of training, can be of immediate practical benefit.

In conclusion, the present study showed the benefits of two different types of training on mental addition problems in children in 3rd and 5th grades: the strategic training was more useful for the younger children, with near transfer effects on subtraction problems and math fluency, while the 5th-graders benefited more from a process-based training, improving their response times in the criterion task.

References

Baroody, A. J., & Dowker, A. (2003). *The development of arithmetic concepts and skills.* Mahwah, NJ: Lawrence Erlbaum Associates.

Bjork, R. A. (1994). Memory and metamemory considerations in the training of human beings. In J. Metcalfe, & A. Shimamura (Eds.), *Metacognition: Knowing about knowing* (pp. 185–205). Cambridge, MA: MIT Press.

Blayney, P., Kalyuga, S., & Sweller, J. (2010). Interactions between the isolated–interactive elements effect and levels of learner expertise: Experimental evidence from an accountancy class. *Instructional Science*, *38*(3), 277–287. http://dx.doi.org/10.1007/s11251-009-9105-x.

Borella, E., Carretti, B., Riboldi, F., & De Beni, R. (2010). Working memory training in older adults: Evidence of transfer and maintenance effects. *Psychology and Aging*, *25*, 767–778. http://dx.doi.org/10.1037/a0020683.

Brunstein, A., Betts, S., & Anderson, J. R. (2009). Practice enables successful learning under minimal guidance. *Journal of Educational Psychology*, *101*, 790–802. http://dx.doi.org/10.1037/a0016656.

Butterworth, B. (1999). *The mathematical brain.* London: Macmillan.

Butterworth, B. (2005). The development of arithmetical abilities. *Journal of Child Psychology and Psychiatry*, *46*, 3–18. http://dx.doi.org/10.1111/j.1469-7610.2004.00374.x.

Butterworth, B., & Laurillard, D. (2010). Low numeracy and dyscalculia: Identification and intervention. *ZDM Mathematics Education*, *42*(6), 527–539. http://dx.doi.org/10.1007/s11858-010-0267-4.

Carpenter, T. P., Franke, M. L., Jacobs, V. R., Fennema, E., & Empson, S. B. (1998). A longitudinal study of invention and understanding in children's multidigit addition and subtraction. *Journal for Research in Mathematics Education*, *29*, 3–20.

Carretti, B., Borella, E., & De Beni, R. (2007). Does a strategic memory training improve the working memory capacity of younger and older adults? *Experimental Psychology*, *54*, 311–320.

Carroll, W. M. (1994). Using worked examples as an instructional support in the algebra classroom. *Journal of Educational Psychology*, *86*, 360–367.

Caviola, S., Mammarella, I. C., Cornoldi, C., & Lucangeli, D. (2009). A metacognitive visuospatial working memory training for children. *International Electronic Journal of Elementary Education*, *2*, 122–136.

Caviola, S., Gerotto, G., Lucangeli, D., & Mammarella, I. C. (2016). *Nuove prove di fluenza per la matematica. [Math fluency tasks].* Trento: Erickson (in press).

Clements, D. H., & Sarama, J. (2007). Effects of a preschool mathematics curriculum: Summative research on the building blocks project. *Journal for Research in Mathematics Education*, *38*, 136–163.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences.* Hillsdale, NJ, US: Lawrence Erlbaum.

Cooper, G. A., & Sweller, J. (1987). The effects of schema acquisition and rule automation on mathematical problem solving transfer. *Journal of Educational Psychology*, *79*, 347–362.

De Smedt, B., Noël, M. P., Gilmore, C., & Ansari, D. (2013). How do symbolic and non-symbolic numerical magnitude processing skills relate to individual differences in children's mathematical skills? A review of evidence from brain and behavior. *Trends in Neuroscience and Education*, *2*, 48–55. http://dx.doi.org/10.1016/j.tine.2013.06.001.

Dehaene, S. (2009). Origins of mathematical intuitions: The case of arithmetic. *Annals of the New York Academy of Sciences*, *1156*, 232–259. http://dx.doi.org/10.1111/j.1749-6632.2009.04469.x.

Delazer, M., Ischebeck, A., Domahs, F., Zamarian, L., Koppelstaetter, F., Siedentopf, C. M., ... Felber, S. (2005). Learning by strategies and learning by drill - Evidence from an fMRI study. *NeuroImage*, *25*, 838–849. http://dx.doi.org/10.1016/j.neuroimage.2004.12.009.

Feigenson, L., Dehaene, S., & Spelke, E. (2004). Core systems of number. *Trends in Cognitive Sciences*, *87*, 307–314. http://dx.doi.org/10.1016/j.tics.2004.05.002.

Frank, M. C., & Barner, D. (2012). Representing exact number visually using mental abacus. *Journal of Experimental Psychology: General*, *141*, 134–149. http://dx.doi.org/10.1037/a0024427.

Fuchs, L. S., Compton, D. L., Fuchs, D., Paulsen, K., Bryant, J. D., & Hamlett, C. L. (2005). The prevention, identification, and cognitive determinants of math difficulty. *Journal of Educational Psychology*, *97*, 493–513. http://dx.doi.org/10.1037/0022-0663.97.3.493.

Fuchs, L. S., Seethaler, P. M., Powell, S. R., Fuchs, D., Hamlet, C. L., & Fletcher, J. M. (2008). Effects of preventative tutoring on the mathematical problem solving of third-grade students with math and reading difficulties. *Exceptional Children*, *74*, 155–173.

Geary, D. (2010). Mathematical disabilities: Reflections on cognitive, neuropsychological, and genetic components. *Learning and Individual Differences*, *20*, 130–133. http://dx.doi.org/10.1016/j.lindif.2009.10.008.

Griffin, S., Case, R., & Siegler, R. (1994). Rightstart: Providing the central conceptual prerequisites for first formal learning of arithmetic to students at-risk for school failure. In K. McGilly (Ed.), *Classroom lessons: Integrating cognitive theory and classroom practice* (pp. 24–49). Cambridge, MA: Bradford Books MIT Press.

Grissom, R. J., & Kim, J. J. (2012). *Effect sizes for research: Univariate and multivariate applications.* New York, NY: Routledge.

Halberda, J., Mazzocco, M. M. M., & Feigenson, L. (2008). Individual differences in non-verbal number acuity correlate with maths achievement. *Nature*, *455*, 665–669. http://dx.doi.org/10.1038/nature07246.

Hedges, L. V., & Olkin, I. (1985). *Statistical methods for meta-analysis.* Orlando, FL: Academic Press.

Jaeggi, S. M., Buschkuehl, M., Jonides, J., & Perrig, W. J. (2008). Improving fluid intelligence with training on working memory. *Proceedings of the National Academy of Sciences, USA. 105*. (pp. 6829–6833). http://dx.doi.org/10.1073/pnas.0801268105.

Kalyuga, S. (2007). Expertise reversal effect and its implications for learner-tailored instruction. *Educational Psychology Review*, *19*, 509–539. http://dx.doi.org/10.1007/s10648-007-9054-3.

Kalyuga, S., Chandler, P., Tuovinen, J., & Sweller, J. (2001). When problem solving is superior to studying worked examples. *Journal of Educational Psychology*, *93*, 579–588. http://dx.doi.org/10.1037/0022-0663.93.3.579.

Kalyuga, S., Ayres, P., Chandler, P., & Sweller, J. (2003). The expertise reversal effect. *Educational Psychology*, *38*, 23–31. http://dx.doi.org/10.1207/S15326985EP3801_4.

Kalyuga, S., Law, Y. K., & Lee, C. H. (2013). Expertise reversal effect in reading Chinese texts with added causal words. *Instructional Science*, *41*(3), 481–497. http://dx.doi.org/10.1007/s11251-012-9239-0.

Karbach, J., Strobach, T., & Schubert, T. (2015). Adaptive working-memory training benefits reading, but not mathematics in middle childhood. *Child Neuropsychology*, *21*(3), 285–301. http://dx.doi.org/10.1080/09297049.2014.899336.

Kucian, K., Grond, U., Rotzer, S., Henzi, B., Schönmann, C., Plangger, F., et al. (2011). Mental number line training in children with developmental dyscalculia. *NeuroImage*, *57*, 782–795. http://dx.doi.org/10.1016/j.neuroimage.2011.01.070.

Lee, H. S., & Anderson, J. R. (2013). Student learning: What has instruction got to do with it? *Annual Review of Psychology*, *64*, 445–469. http://dx.doi.org/10.1146/annurev-psych-113011-143833.

Lee, C. H., & Kalyuga, S. (2014). Expertise reversal effect and its instructional implications. *Acknowledgments and Dedication* (pp. 31–44).

LeFevre, J. -A., DeStefano, D., Penner-Wilger, M., & Daley, K. (2006). Selection of procedures in mental subtraction. *Canadian Journal of Experimental Psychology*, *60*, 209–220. http://dx.doi.org/10.1037/cjep2006020.

Lewis, C., Hitch, G. J., & Walker, P. (1994). The prevalence of specific arithmetic difficulties and specific reading difficulties in 9- to 10-year-old boys and girls. *Journal of Child Psychology and Psychiatry*, *35*, 283–292. http://dx.doi.org/10.1111/j.1469-7610.1994.tb01162.x.

Lustig, C., & Flegal, K. E. (2008). Targeting latent function: Encouraging effective encoding for successful memory training and transfer. *Psychology and Aging*, *23*(4), 754–764. http://dx.doi.org/10.1037/a0014295.

Mayer, R. E., & Sims, V. K. (1994). For whom is a picture worth a thousand words? Extensions of a dual-coding theory of multimedia learning. *Journal of Educational Psychology*, *84*, 389–460.

Obersteiner, A., Reiss, K., & Ufer, S. (2013). How training on exact or approximate mental representations of number can enhance first-grade students' basic number processing and arithmetic skills. *Learning and Individual Differences*, *23*, 125–135. http://dx.doi.org/10.1016/j.learninstruc.2012.08.004.

Park, D. C., Gutchess, A. H., Meade, M. L., & Stine-Morrow, E. A. (2007). Improving cognitive function in older adults: Nontraditional approaches. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, *62*(1), 45–52.

Prentice, D. A., & Miller, D. T. (1992a). *When small effects are impressive.*

Prentice, D. A., & Miller, D. T. (1992b). *When small effects are impressive.*

Prentice, D. A., & Miller, D. T. (1992c). When small effects are impressive. *Psychological Bulletin*, *112*(1), 160–164.

Räsänen, P. (2015). Computer-assisted interventions on basic number skills. In R. Cohen Kadosh, & A. Dowker (Eds.), *The Oxford handbook of mathematical cognition* (pp. 745–766). Oxford, UK: Oxford University Press.

Räsänen, P., Salminen, J., Wilson, A. J., Aunio, P., & Dehaene, S. (2009). Computer-assisted intervention for children with low numeracy skills. *Cognitive Development*, *24*, 450–472. http://dx.doi.org/10.1016/j.cogdev.2009.09.003.

Renkl, A. (1997). Learning from worked-out examples: A study on individual differences. *Cognitive Science*, *21*, 1–29. http://dx.doi.org/10.1207/s15516709cog2101_1.

Sasanguie, D., Göbel, S. M., Moll, K., Smets, K., & Reynvoet, B. (2013). Approximate number sense, symbolic number processing, or number-space mappings: What underlies mathematics achievement? *Journal of Experimental Child Psychology*, *114*, 418–431. http://dx.doi.org/10.1016/j.jecp.2012.10.012.

Schmidt, R. A., & Bjork, R. A. (1992). New conceptualization of practice: Common principles in three paradigms suggest new concepts for training. *Psychological Science*, *3*, 207–217. http://dx.doi.org/10.1111/j.1467-9280.1992.tb00029.x.

Schubert, T., Strobach, T., & Karbach, J. (2014). New directions in cognitive training: On methods, transfer, and application. *Psychological Research*, *78*(6), 749–755. http://dx.doi.org/10.1007/s00426-014-0619-8.

Shalev, R. S., & Gross-Tsur, V. (2001). Developmental dyscalculia. *Pediatric Neurology*, *24*, 337–342. http://dx.doi.org/10.1016/S0887-8994(00)00258-7.

Stukas, A. A., & Cumming, G. (2014). Interpreting effect sizes: Toward a quantitative cumulative social psychology. *European Journal of Social Psychology*, *44*(7), 711–722. http://dx.doi.org/10.1002/ejsp.2019.

Swanson, H. L. (2011). Working memory, attention, and mathematical problem solving: A longitudinal study of elementary school children. *Journal of Educational Psychology*, *103*, 821–837. http://dx.doi.org/10.1037/a0025114.

Sweller, J., & Cooper, G. A. (1985). The use of worked examples as a substitute for problem solving in learning algebra. *Cognition and Instruction*, *2*, 59–89.

Thurstone, L. L., & Thurstone, T. G. (1963). *Primary mental abilities.* Chicago, IL: Science Research.

Titz, C., & Karbach, J. (2014). Working memory and executive functions: Effects of training on academic achievement. *Psychological Research*, *78*(6), 852–868. http://dx.doi.org/10.1007/s00426-013-0537-1.

Toll, S. W. M., & Van Luit, J. E. H. (2014). Effects of remedial numeracy instruction throughout kindergarten starting at different ages: Evidence from a large-scale longitudinal study. *Learning and Instruction*, *25*, 45–54. http://dx.doi.org/10.1016/j.lindif.2013.03.006.

Verhaeghen, P., Marcoen, A., & Goossens, L. (1992). Improving memory performance in the aged through mnemonic training: A meta-analytic study. *Psychology and Aging*, *7*(2), 242–251.

Wilson, A. J., Dehaene, S., Pinel, P., Revkin, S. K., Cohen, L., & Cohen, D. (2006). Principles underlying the design of ≪The Number Race≫, an adaptive computer game for remediation of dyscalculia. *Behavioral and Brain Functions*, *2*, 19. http://dx.doi.org/10.1186/1744-9081-2-19.

Wilson, A. J., Dehaene, S., Dubois, O., & Fayol, M. (2009). Effects of an adaptive game intervention on accessing number sense in low-socioeconomic-status kindergarten children. *Mind, Brain, and Education*, *3*, 223–224. http://dx.doi.org/10.1111/j.1751-228X.2009.01075.x.