Learning and Instruction 46 (2016) 1–11

Contents lists available at ScienceDirect

Learning and Instruction

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

Enhancing arithmetic in pre-schoolers with comparison or number line estimation training: Does it matter?☆

Bieke Maertens [a, b], Bert De Smedt [c], Delphine Sasanguie [a, b], Jan Elen [d], Bert Reynvoet [a, b, *]

[a] Brain & Cognition, KU Leuven, Leuven, Belgium
[b] Faculty of Psychology and Educational Sciences@ Kulak, Kortrijk, Belgium
[c] Parenting and Special Education Research Group, KU Leuven, Leuven, Belgium
[d] Centre for Instructional Psychology and Technology, KU Leuven, Leuven, Belgium

ARTICLE INFO

Article history:
Received 24 June 2015
Received in revised form
15 July 2016
Accepted 4 August 2016
Available online 7 September 2016

Keywords:
Magnitude processing training
Numerical intervention
Arithmetic achievement

ABSTRACT

Children's intuitive understanding of number, i.e. number sense, is associated with individual differences in mathematics achievement. To investigate the causal association between number sense, traditionally assessed with comparison or number line estimation tasks, and mathematics achievement, often assessed with an arithmetic test, an intervention study was conducted that aimed at training either comparison or number line estimation skills. We contrasted a comparison and number line estimation training. By doing so, we wanted to address the question which intervention had the largest effect on arithmetic. In addition, such a direct comparison between comparison and number line estimation trainings would allow us to get more insight in the association between both tasks. Participants were 151 five-year-olds that were randomly allocated to either an experimental condition (i.e. comparison or number line estimation) or one of the two control conditions (i.e., active control condition and empty control condition) in a pretest-posttest design measuring number knowledge, (non-)symbolic comparison and number line estimation and arithmetic. The results showed that both comparison and number line estimation trainings had a positive effect on arithmetic. However, the absence of transfer effects from one task to another, also suggested that comparison and number line estimation rely on different mechanisms and probably influence arithmetic through different mechanisms.

1. Introduction

Mathematical ability plays an important role in one's daily-life and professional activities. For instance, Dieckman, Slovic, and Peters (2009) observed that people low in mathematical ability have problems with integrating numerical information in decision making. Although mathematical ability is also influenced by environmental factors like socioeconomic status (e.g., Jordan & Levine, 2009) and general cognitive factors like working memory (e.g., Friso-van den Bos, van der Ven, Kroesbergen, & van Luit, 2013), there is now considerable evidence to suggest that mathematical ability also builds upon an intuitive understanding of number, i.e. number sense (Dehaene, 1997) that can already be observed in very young children through the presence of a set of basic number skills (e.g., Halberda & Feigenson, 2008; Holloway & Ansari, 2009; Sasanguie, De Smedt, Defever, & Reynvoet, 2012). This observation has motivated researchers to develop training programs that focused on (some of) these basic number skills in order to improve children's mathematical achievement (e.g., Ramani & Siegler, 2008; Wilson, Dehaene, Dubois, & Fayol, 2009). However, it is still unclear whether these different numerical training programs are equally effective. In this study, we compared the effects of two different interventions, training number comparison and on number line estimation skills. Both skills are assumed to operate on the representation of number (Laski & Siegler, 2007). By analysing the effects of both interventions on other number skills and arithmetic, we will examine whether both skills build on a common underlying mechanism and clarify which intervention has the largest effect on arithmetic ability.

☆ This research was supported by grant GOA 2012/010 of the Research Fund KU Leuven, Belgium. Delphine Sasanguie is a post-doctoral reseracher of the Fund for Scientific Research - Flanders. We would like to thank all participating children and teachers.

* Corresponding author. Brain & Cognition, KU Leuven, Faculty of Psychology and Educational Sciences @ Kulak, Etienne Sabbelaan 53, 8500 Kortrijk, Belgium.
E-mail address: bert.reynvoet@kuleuven-kulak.be (B. Reynvoet).

http://dx.doi.org/10.1016/j.learninstruc.2016.08.004
0959-4752/

1.1. Number sense

Researchers have defined the term number sense in different ways. For instance, Berch (2005) made a list of more than 30 definitions ranging from elementary definitions about quantity, a typical definition from cognitive psychology, to more complex skills, such as flexibility with operations and procedures, a definition shared among mathematics educators (see also Reynvoet, Smets, & Sasanguie, 2016). Here, we define number sense as is typically done in cognitive psychology. Number sense is considered as the intuitive understanding of numbers that is demonstrated through successful discrimination of visual sets on the basis of number or understanding the result of manipulations (e.g., adding or subtracting elements) with such visual sets (Dehaene, 1997). These abilities are grounded in a cognitive system often referred to as the 'approximate number system' (ANS). This system makes it possible for infants, but also non-human species to discriminate non-symbolic numbers (e.g., Nieder & Dehaene, 2009). Several authors assume that, later in development, numerical symbols (e.g., Arabic number symbols) are mapped on this ANS (e.g., Barth, La Mont, Lipton, & Spelke, 2005; Holloway & Ansari, 2009; Mundy & Gilmore, 2009, but see; Carey, 2009). As a consequence, the concept number sense implicitly also refers to the representation of numerical symbols because they activate the same representation as a numerical equivalent visual set of objects.

The characteristics of this number representation are typically assessed with a number comparison task. In a comparison task, participants have to indicate which of two presented stimuli (i.e., digits or dot arrays) is more. Performance on this task results in a distance effect and a size effect (e.g. Holloway & Ansari, 2009). The distance effect refers to the observation that decisions are more difficult when the numerical distance between the stimuli is small. The size effect reflects more difficult discriminations for numerical larger numbers. Both observations resulted in the popular idea that numbers are represented on a 'mental number line', which obeys Weber-Fechner's law and represents numbers in an approximate and compressed way (i.e. noisier representations for larger numbers, Dehaene, 1997; Gallistel & Gelman, 1992).

The mental number line concept has inspired Siegler and Booth (Booth & Siegler, 2006; Siegler & Booth, 2004) to assess number representations through number line estimation tasks, an approach that was followed by many others (e.g. Dehaene, Izard, Spelke, & Pica, 2008; Link, Nuerk, & Moeller, 2014). In the number line estimation task, participants have to indicate the position of a non-symbolic or symbolic number on an empty line. The idea behind this task is that the positioning of numbers on the external line reflects the organisation of the underlying mental number line.

However, the idea that number comparison and number line estimation skills rely on the same representation of numbers has been recently debated (e.g. Barth & Paladino, 2011; Link et al., 2014; Sasanguie & Reynvoet, 2013). For instance, Sasanguie and Reynvoet (2013) examined the association between the size effect observed in the comparison task and the amount of compression for larger numbers observed in the number line estimation task. The association between both indices was not significant, suggesting that different representations or mechanisms might be stake. A possible explanation for this unrelated performance can be found in the study of Barth and Paladino (2011). They argued that subjects use anchor points to place a number on an empty number line. As a consequence, number line placements are based on proportion reasoning, taking the visual begin and endpoints and, at a later age even the implicit midpoints of a line, into account. This requires spatial reasoning as well as arithmetic skills to perform this task, which is less important in the comparison task (see also Link et al., 2014).

1.2. Association between basic number skills and mathematical achievement

Although it is debated whether the same representation is involved in both tasks, many studies have shown that the performance on both tasks is related to (later) individual differences in mathematical achievement. Many studies have revealed a predictive association between number comparison and mathematical achievement: fast and/or accurate comparison of numbers is related to better performance on a general mathematical achievement test (e.g., De Smedt, Verschaffel, & Ghesquière, 2009; Feigenson, Libertus, & Halberda, 2013; Sasanguie, Göbel, Moll, Smets, & Reynvoet, 2013). Similarly, the accurate positioning of the numbers in a number line estimation task is related to better performance on a general mathematical achievement test (Booth & Siegler, 2006; Link et al., 2014; Sasanguie et al., 2013).

Currently, researchers still disagree whether non-symbolic or symbolic number processing is more related to later mathematical achievement. Several studies have shown that discrimination of non-symbolic stimuli is related to the score on a mathematical achievement test (e.g. Feigenson et al., 2013; Halberda, Mazzocco, & Feigenson, 2008; Piazza et al., 2010). In contrast, other studies found that the processing of symbolic numbers, but not non-symbolic numbers, is related to mathematical achievement (e.g. Holloway & Ansari, 2009; Sasanguie et al., 2012; Vanbinst, Ghesquière, & De Smedt, 2015). To date, these different positions concerning what is *the best* predictor for mathematics achievement, symbolic or non-symbolic processing, still exist. However, narrative reviews (e.g. De Smedt, Noël, Gilmore, & Ansari, 2013) and a meta-analysis (Schneider et al., 2016) indicate that, although both are related to mathematics achievement, the relation between symbolic number processing and mathematics achievement is significantly more robust.

1.3. Training studies

To investigate whether number sense and mathematical achievement are causally related, several training studies were conducted to investigate whether mathematical achievement, in several studies assessed with an arithmetic test, improved after training number sense abilities such as comparison or number line estimation (for a review, see De Smedt et al., 2013). For example, some intervention studies have made use of *"The Number Race"* game (Wilson et al., 2006), a computer program in which children were presented different formats of comparison tasks, adapted to their performance level (Obersteiner, Reiss, & Ufer, 2013; Räsänen, Salminen, Wilson, Aunio, & Dehaene, 2009; Wilson et al., 2009). Although the game contained a lot of practice with comparison, children also had to count while moving an object on a number board when they answered correctly. After training, children performed more accurate and faster on comparison tasks and in some of the studies, also improvement in arithmetic performance was observed (Obersteiner et al., 2013; Wilson et al., 2009). In another study, Hyde, Khanum, and Spelke (2014) trained first graders either with a non-symbolic comparison task or with a non-symbolic addition task, in which two non-symbolic stimuli had to be added first and then compared with a third stimulus. Both trainings led to improvement in symbolic arithmetic. Park and Brannon (2014) showed that also adults benefit from practicing with non-symbolic addition: Non-symbolic addition training resulted in improvement on two and three digit addition problems, whereas practice with other tasks did not.

Other training studies have focused on number line estimation (e.g., Kucian et al., 2011; Link, Moeller, Huber, Fischer, &

Nuerk, 2013; Ramani & Siegler, 2008, 2011; Siegler & Ramani, 2009). For instance, in a series of studies, Ramani and Siegler embedded this task into a board game in which children had to move a token a certain number of squares and investigated which type of board game (i.e., linear versus circular number board game or number versus coloured board game) induced the largest effect on children's number line estimation. Playing with the linear number board game induced larger improvements on number line estimation compared to the circular one (Ramani & Siegler, 2011; Siegler & Ramani, 2009) or the coloured version (Ramani & Siegler, 2008; Ramani, Siegler, & Hitti, 2012). These authors also observed transfer effects to arithmetic after playing with the linear number board version (Ramani & Siegler, 2011; Siegler & Ramani, 2009). Kucian et al. (2011) developed a computerized game, *"Rescue Calcularis"*, in which children had to position the magnitude or the outcome of the arithmetic problem on the number line. Results showed enhanced performance on number line estimation tasks and improvement on arithmetic.

1.4. The present study

The goal of the present study was to compare the effects of comparison and number line training on mathematical achievement. Previous studies have shown that training of both tasks resulted in higher mathematical achievement, demonstrating the causal role of both comparison and number line estimation skills. However, the effects of both type of trainings on mathematical achievement have not been compared directly, leaving the question what the most effective training is, unanswered. A first goal of the present study was to provide an answer on this outstanding question. The second goal of this study was to further investigate the relation between comparison and number line performance. As we have mentioned, it is debated whether both tasks rely on the same representation of numbers, i.e. mental number line. We will do this by investigating the effect of a specific number sense training (comparison vs. number line training) on comparison and number line estimation skills. If comparison and number line tasks reflect the same underlying mental number line, we expected that practising comparison would generalize to improvements in number line estimation skills and vice versa. In contrast, if both skills rely on different mechanisms, no transfer effects to the other, untrained number sense skill, are expected.

To investigate these questions, we used the *K-games* of *Dudeman & Sidegirl: Operation clean world* (Linsen et al., 2015). These tablet games were developed for children in the final year of kindergarten in Flanders (Belgium) (i.e. around the age of five and a half). Kindergarten children were chosen because children at this age are still in the process of acquiring the meaning of Arabic digits (Vlaams Ministerie van Onderwijs en Vorming Agentschap voor Kwaliteitszorg in Onderwijs en Vorming, 2010), a process that is explicitly targeted by these games. An active and an empty control condition were included to assure that the improvements on the dependent measures were due to number sense training. In the active control condition, children had to play with *Mega memory Match*, an existing game on tablets (GiggleUp Kids Apps And Educational Games Pty Ltd, 2013). This active control condition was included to control for effects of motivation - playing with a tablet is an activity the kindergartners typically do not do during class. In the empty control group, children did not play with a tablet game and stayed in their classroom participating in the normal classroom activities. This empty control condition was included to control for test-retest effects.

2. Method

2.1. Participants

Originally, 160 children in the third year of kindergarten[1] from ten classes from five different schools participated in the study. Most of these children were from middle-to-high socio-economic status (SES) families as indicated by the highest educational degree of the mothers that was secondary education (31%), bachelor level (34%) or master level (30%). Five percent did not respond to this question. Five children (2 children in the number line estimation condition, 2 children in the comparison condition, and 1 child in the active control condition) did not have Dutch as their mother tongue, but their level of Dutch was sufficient to attend the class and the intervention. Within each class, children were randomly assigned to one of the four intervention conditions. Four children (2 children in the number line estimation condition, and 2 children in the comparison condition) were excluded from analyses because they were frequently absent during the intervention period. Another five children (3 children in the number line estimation condition, 1 child in the comparison condition, and 1 child in the active control condition) were also excluded because they were absent at the post-test. The final sample thus consisted of 151 kindergartners, of which 47 children participated in the comparison condition (M_{age} (t_1) = 5.44 years, SD = 0.31; 23 boys), 41 children in the number line estimation condition (M_{age} (t_1) = 5.49 years, SD = 0.29; 18 boys), 37 children in the active control condition (M_{age} (t_1) = 5.43 years, SD = 0.27; 14 boys), and 26 children in the empty control condition (M_{age} (t_1) = 5.32 years, SD = 0.31; 14 boys).

2.2. Procedure and materials

The intervention study was conducted in the first quarter of the school year (September–November). In the pre- and post-test, children's number sense and arithmetic were measured. During the intervention, children in the comparison, number line estimation, and active control condition played 6 sessions of approximately 10 min with a tablet game over a period of three weeks, resulting in a total practice time of about 60 min. These children were trained in a quiet separate room in the school, in groups of about five children. In each session, they were instructed to complete two or three levels. After completing the levels in each session, the child had to log out from the game. In the following session, the game restarted automatically at the level where the player ended the previous time. Children wore headphones so they could hear the auditory instructions and feedback. An experimenter was present during all sessions to log-on the children on the tablet and to help them with possible practical problems. The children in the empty control condition did not play with a tablet game, and had to stay in the classroom and participated in regular kindergarten activities.

2.3. Pre- and post-measures

In the pre- and posttest, two number knowledge tasks were first included to investigate whether children were familiar with Arabic digits and their semantic meaning. Children's number sense skills

[1] In Belgium, compulsory education starts in primary school (i.e. around the age of 6 years). However, although it is not compulsory, the majority of children typically start in kindergarten at about 3 years old. Kindergarten consists of 3 grades. In the present sample, all children started school in the first year of kindergarten and they were now in their third year of kindergarten.

were measured with two comparison tasks and two number line estimation tasks. These tasks were run on a tablet (iPad 2 Wi-Fi 16 GB with 9.7 inches display). Screenshots of number knowledge, comparison and number line estimation tasks are shown in Fig. 1. In addition to these tasks, children also completed two standardized arithmetic tests.

2.3.1. Number knowledge tasks

In a variant of the *give a number* task (Condry & Spelke, 2008), children were shown the digits 3, 5, 7, and 9 in random order and for each digit they had to tap the same number of dots on the tablet screen.

In the *connecting task*, an Arabic digit was shown and the child had to select which of three collections of dots corresponded to that digit. Incorrect response alternatives differed in numerical distance from the correct response, one alternative with one digit away and one with more than two digits away. All single digits were included in this task. For each correct answer, the child was given one point.

2.3.2. Comparison tasks

The stimuli in the symbolic comparison task were all single digits; 16 trials had a numerical distance of 1 and 16 trials with a numerical distance of 4, resulting in 32 trials in total. The stimuli in the non-symbolic tasks were dot arrays. One of the dot arrays always contained the reference numerosity 16, whereas the other dot array contained 8, 11, 13, 19, 24, and 32 dots. Sixteen trials of each of

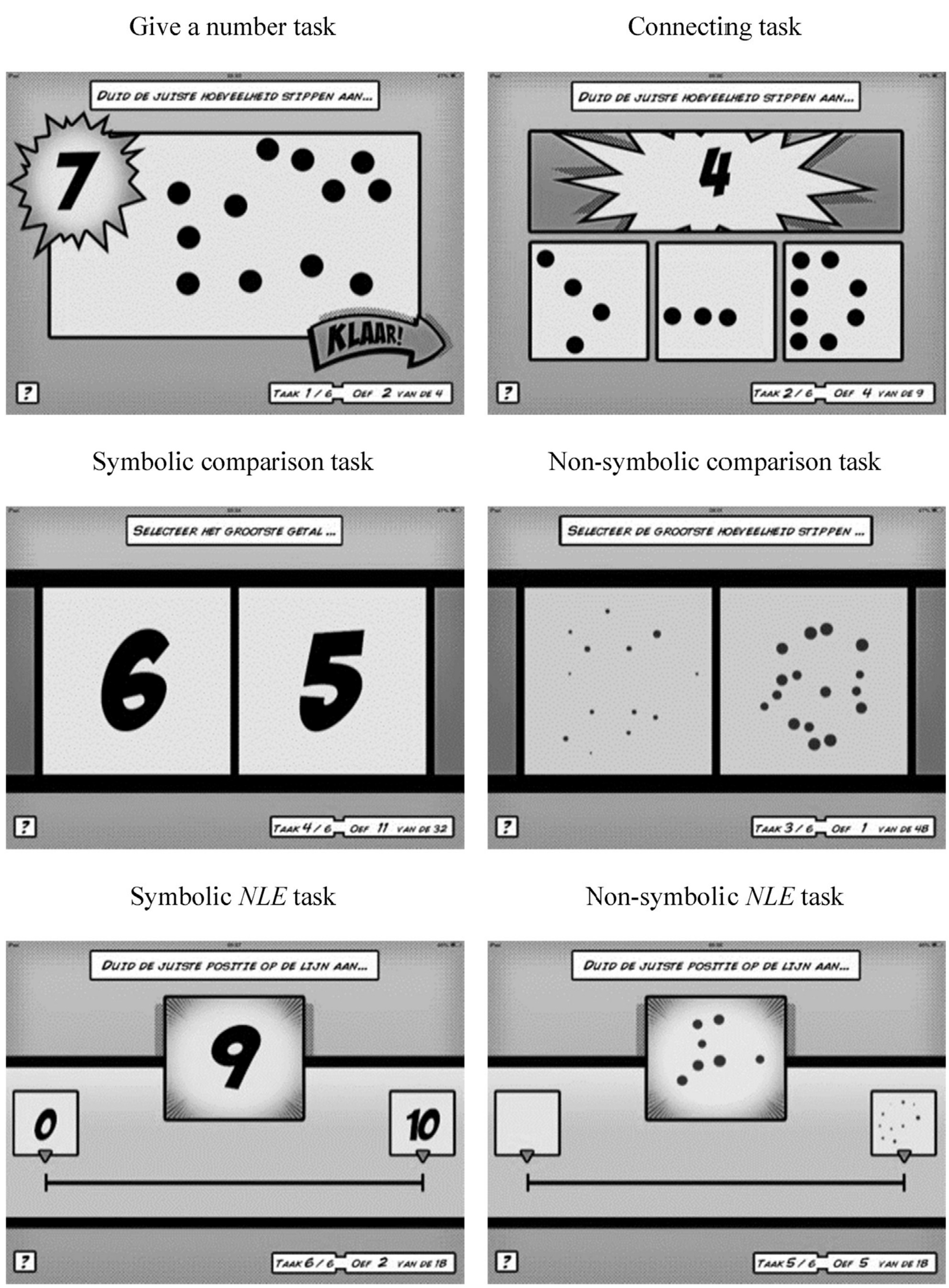

Fig. 1. Screenshots of the number knowledge, comparison and number line estimation tasks.

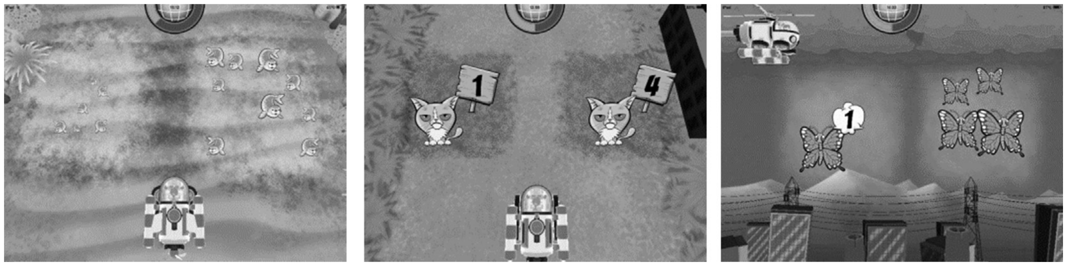

Fig. 2. Screenshots of some levels in the *K* − *comparison* game.

the ratios (2.00, 1.50, and 1.20) were included, resulting in a total of 48 trials. These stimuli were generated with the MatLab script developed by Gebuis and Reynvoet (2011), controlling for four visual parameters (i.e., convex hull, total surface, item size, and density). Regression analyses showed that differences in the visual parameters did not explain the variance in numerical distance (all $R^2 < 0.05$, $ps > 0.17$).

The procedure in both comparison tasks was the same. Two stimuli were presented simultaneously on the left and right side of the tablet. The stimuli were presented for 1500 ms, followed by a blank screen. The children had to answer during stimulus display or during the blank screen by tapping on the side of the numerical larger stimulus. After responding, an interstimulus interval of 600 ms followed after which the next trial started. The child was instructed to answer as accurate and as fast as possible, though only accuracy was analysed as reaction times collected on different tablets may be unreliable due to timing errors (Schatz, Ybarra, & Leitner, 2015). There were three practice trials to become familiar with the task demands.

2.3.3. *Number line estimation tasks*

In the number line estimation tasks, children had to place a non-symbolic or symbolic numbers on an empty number line. A line of 14 cm was presented in the middle of the tablet labelled by "0" (symbolic variant) or an empty circle (non-symbolic variant) at the left end point and by "10" (symbolic variant) or a circle with 10 dots (non-symbolic variant) on the right side. The to be positioned number was shown in the middle of the screen, 2.2 cm above the number line. All numbers from 1 to 9 were shown in a random order and had to be positioned twice on the number line in both symbolic and non-symbolic tasks, resulting in a total of 18 trials in each task. Children were instructed to answer as accurately as possible. Three practice trials preceded the task administration.

2.3.4. *Arithmetic measures*

Children's arithmetic skills were measured with two subtests of the TEDI-MATH (Grégoire, Noël, & Van Nieuwenhoven, 2004). The first subtest comprised six pictorially presented *arithmetic problems*, which were read aloud to the child (e.g. "Here you see two red balloons and three blue balloons. How many balloons are there together?"). For each correct answer, the child was given one point. The second subtest comprised 18 horizontally presented *symbolic addition operations* (e.g. "6 + 3 = …"). Only the first problem was read aloud by the experimenter, after which the child had to complete the remaining problems. The child had to solve as many addition problems as possible and testing was stopped after five consecutive errors. The total score was the number of correctly answered problems.

2.4. *Intervention conditions*

2.4.1. *Experimental conditions: comparison and number line estimation*

In both conditions, children were presented with the story that the planet was destroyed and the child needed to help the game characters to make the planet beautiful again by collecting the animals that were hidden in the water, on the land, and in the air. The different levels of the game were characterised by increasing difficulty. The children had to reach a minimum mean accuracy in a level (i.e., 80% in the comparison game and 50% in the number line game) before going to the next level. As long as they did not reach that score, they had to repeat the level. In the comparison training, the difficulty of the levels was based on the numbers (i.e. 1–4, 1–9 and 5–18), the display duration (i.e. infinite or 1500 ms) and the type of stimuli (i.e. non-symbolic, symbolic or mixed). In the number line training, level difficulty was based on the number of anchor points (i.e. 9, 1 or 0), display duration of the stimulus (i.e. infinite or 1500 ms) and type of stimuli (i.e. non-symbolic, symbolic or mixed). As a consequence some children played somewhat longer than others, but importantly, all children finished the last level and thereby received the same learning content. In that way, all children trained at their individual performance level and at their own pace. After each session, the children received pictures of the animals they had caught as a reward. The two games used in these conditions are described into detail in Linsen et al. (2015), where they are denoted as K-games as they were designed for kindergartners.

In the *comparison game*, children had to navigate their vehicle through the world and they were shown two stimuli on the left and right side of the tablet screen (Fig. 2). They were instructed to collect as many animals as possible by tapping on the numerically larger stimuli. Immediately after tapping, the vehicle moved to that position and if the answer was correct, the animals were collected. The feedback was extended by a positive or negative sound depending on whether the answer was correct or wrong. After each session, the children received printed pictures of the collected animals as reward.

In the *number line game*, children had to navigate their vehicle and were shown an empty number line on which they had to position a number (Fig. 3). Here too, they were instructed to collect as many animals as possible by tapping on corresponding position on the number line of the presented numbers. Immediately after tapping on that position, the vehicle moved over there and grasped for the animal. If the answer was correct, a hidden animal came out of the line and disappeared in the vehicle. If the answer was incorrect, the animal came out, but stayed at its position. This feedback was augmented with a positive (high tone) or negative (low tone) feedback. Here too, the children received pictures of the animals as reward after each session.

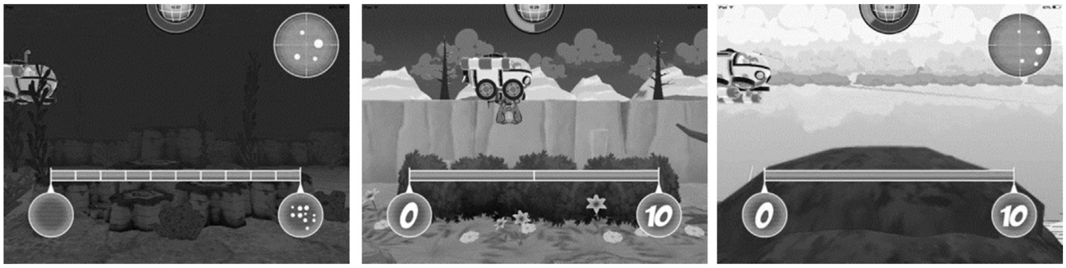

Fig. 3. Screenshots of some levels in the $K - number\ line$ game.

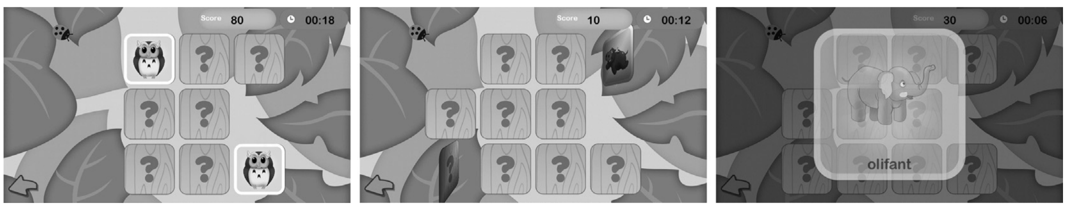

Fig. 4. Screenshots of the *Mega Memory Match*-game, used in the active control condition.

2.4.2. Control conditions: active and empty control condition

In the *active control condition* children played an existing, commercial game called 'Mega Memory Match' on a tablet (Fig. 4; GiggleUp Kids Apps And Educational Games Pty Ltd, 2013). This is a typical memory game with different topics (e.g., animals, vehicles, food, letters…) but no numbers. Children had to play three different levels each session. This approximately matched the playtime of the children in the experimental conditions. In each session, we gradually adjusted the degree of difficulty by increasing the number of cards that comprised the memory game, going from a 2×2 to a 7×4 grid and by decreasing the amount of time the picture side is shown in the beginning (from 10 s to 0 s). As in the other experimental conditions, children received auditory and visual feedback. For a correct match of cards, a positive sound was played and the cards were collected. For an incorrect match, a bummer sound was played and the cards remained on their position. During the entire intervention, the children heard soft, happy background music. After each session, the children received pictures of animals as a reward.

The children in the *empty control condition* did not play with a tablet game. They had to stay in the classroom doing activities their teacher instructed them to do (e.g., doing arts and crafts, listening to a story, or playing with dolls).

3. Results

3.1. Descriptive statistics

Analyses showed that age, $F(3,147) = 1.77$, $p = 0.156$, $\eta_p^2 = 0.036$, and gender,[2] $\chi^2(3) = 1.87$, $p = 0.600$ did not statistically differ between conditions. Table 1 shows the mean performance on all tasks during pre- and post-test in all conditions, together with the corresponding gain scores.

Table 1
Performance at pretest and posttest and gain scores in all conditions.

Measure	Pretest		Posttest		Gain scores	
	M	*SD*	*M*	*SD*	*M*	SD
Give a number (*% correct*)						
Comparison	52.13	35.67	63.83	30.31	11.70	37.54
NLE	45.73	36.19	73.17	29.26	27.44	38.65
Active control	54.73	38.11	64.19	35.13	9.46	41.81
Empty control	46.15	37.21	60.58	36.86	14.42	36.86
Connecting (*% correct*)						
Comparison	68.79	21.44	81.09	18.67	12.29	27.53
NLE	70.46	28.18	78.05	22.01	7.59	27.49
Active control	75.38	21.46	73.57	20.01	-0.80	19.16
Empty control	73.08	24.78	74.79	25.64	1.71	26.98
Non-symbolic comparison (*% correct*)						
Comparison	58.87	11.28	63.96	8.92	5.10	11.21
NLE	61.18	9.19	60.52	9.94	−0.66	11.59
Active control	58.67	8.63	61.09	11.89	2.42	9.84
Empty control	58.25	9.97	60.74	10.22	2.48	11.10
Symbolic comparison (*% correct*)						
Comparison	66.49	16.27	73.6	13.46	7.11	15.64
NLE	67.61	16.37	70.43	17.69	2.82	14.27
Active control	64.27	14.63	64.27	18.16	0.00	12.99
Empty control	65.14	14.70	67.79	17.07	2.64	15.48
Non-symbolic NLE (*PAE*)						
Comparison	27.65	10.21	29.48	10.54	1.83	8.11
NLE	29.20	7.51	19.29	9.25	−9.91	9.44
Active control	28.48	8.06	26.07	7.96	−2.41	9.01
Empty control	28.61	7.89	28.64	8.06	0.03	8.57
Symbolic NLE (*PAE*)						
Comparison	26.61	11.38	25.95	10.6	−0.67	7.61
NLE	26.26	11.98	13.99	8.71	−12.27	11.35
Active control	25.39	9.61	24.06	9.91	−1.33	8.83
Empty control	26.56	9.16	25.70	10.36	−0.86	9.31
Arithmetic problems (*# correct*)						
Comparison	3,00	1.74	3.87	1.81	0.87	1.66
NLE	3.10	1.83	4.02	1.60	0.93	1.44
Active control	3.62	1.69	4.19	1.85	0.57	1.41
Empty control	3.54	1.86	3.54	1.92	0.00	1.9
Addition operations (*# correct*)						
Comparison	1.64	2.34	2.19	2.78	0.55	1.59
NLE	1.54	2.13	1.85	2.01	0.32	1.52
Active control	1.30	1.79	1.86	2.39	0.57	1.54
Empty control	1.46	1.77	2.19	2.19	0.73	1.56

Note. Comparison (n = 47); NLE (n = 41); Active control (n = 37); Empty control (n = 26); PAE = percentage absolute error.

[2] On seven (out of eight) of the pretest measures, no gender differences were observed ($ps > 0.16$). On the symbolic addition task boys outperformed girls, $F(1,149) = 2.49$, $p = 0.014$, however, this difference did not survive the Bonferroni correction ($p < 0.05/8 = 0.006$) for multiple comparisons. No gender differences were observed after Bonferroni corrections in the gain scores of the experimental tasks ($ps > 0.06$) indicating that improvements were not moderated by gender.

The mean accuracies on the two number knowledge tasks showed that the children performed on average above chance on this task.

The mean accuracy on the non-symbolic comparison task at pre-test was 59%. Crucially and as expected, performance was influenced by numerical ratio, $F(2,146) = 10.79$, $p < 0.001$, $\eta^2_p = 0.129$, reflecting higher accuracies for larger ratios. The mean accuracy in the symbolic comparison task at pre-test was 66%. Also here, a distance effect was observed in the symbolic comparison task, $F(1,147) = 45.74$, $p < 0.001$, $\eta^2_p = 0.237$. Children were more accurate when the numerical difference between two digits was larger.

In the number line estimation tasks, the percentage absolute error (*PAE*) was used as outcome. The *PAE* was calculated according to the formula of Siegler and Booth (2004): |(estimate − actual number)/scale of estimates|. For example, if a child was asked to estimate 4 on a 0–10 number line and marks the line at the point corresponding to 5.3, the *PAE* would be |(5.3−4)/10| or 13% error. In general, taking into account time of testing during school year (i.e. the present pre-test was conducted in second month of the school year) mean accuracies and variation for comparison and number line estimation tasks were comparable with previous studies with pre-schoolers (e.g. Berteletti, Lucangeli, Piazza, Dehaene, & Zorzi, 2010; Halberda & Feigenson, 2008; Sasanguie et al., 2012; Sasanguie, Defever, Maertens, & Reynvoet, 2014).

In the arithmetic tasks, on average three out of six pictorially presented problems were solved correctly. In contrast, the performance on symbolic addition operation (e.g. "2 + 3 = …") was very low.

One-way ANOVA's on all pre-test measures showed no differences in performance between the four conditions (*ps* > 0.32).

3.2. Reliabilities of tasks and correlations between tasks

The reliability of the number knowledge, comparison, number line estimation and arithmetic tasks was assessed by computing the test-retest reliability (pre and post-test) of these tasks in the two control groups. These analyses indicated low to moderate reliabilities between 0.39 and 0.74 (give a number, $r(63) = 0.42$, $p = 0.001$, connecting, $r(63) = 0.50$, $p < 0.001$, non-symbolic comparison, $r(63) = 0.50$, $p < 0.001$, symbolic comparison, $r(63) = 0.64$, $p < 0.001$, non-symbolic number line, $r(63) = 0.39$, $p = 0.002$, symbolic number line, $r(63) = 0.58$, $p < 0.001$, pictorial arithmetic problems, $r(63) = 0.60$, $p < 0.001$, symbolic addition, $r(63) = 0.74$, $p < 0.001$). These low to moderate reliabilities of the tasks are probably due to the low number of observations in each task, ranging from 4 (give a number task) to 48 (non-symbolic comparison). Lindskog, Winman, Juslin, and Poom (2013) previously demonstrated that the reliability of a numerical task increases when the number of trials increases. Another reason for the low reliabilities is that there was quite some time between pre- and posttest. Substantial development might have taken place in these children, influencing the stability of the skills under investigation. Low reliabilities will add more measurement error making it more difficult to detect differences between conditions or resulting in the underestimation of differences (Hunter & Schmidt, 1990).

The correlations between all tasks are presented in Table 2. These correlations showed that arithmetic performance was significantly related to number knowledge, symbolic comparison and symbolic number line estimation (all *ps* < 0.001). The relation between arithmetic performance and non-symbolic comparison was not significant.

3.3. Effects of training

The training effect was evaluated with a repeated measures ANOVA with *time* (pre-test vs. post-test) as within-subject factor and *condition* (comparison, number line, active control, and empty control) as between-subjects factor on all subtests (Fig. 5). Partial eta-squared was calculated as a measure of effect size. If an interaction effect was found, planned pairwise comparison tests on pre- and post-test scores were conducted correcting for multiple comparisons with a Bonferroni correction ($p < 0.05/8 = 0.006$). Effect sizes (d) were reported for these pairwise comparison using the formula reported in Borenstein (2009).

3.3.1. Number knowledge tasks

On the *give a number task*, a main effect of moment was found, $F(1,147) = 23.71$, $p < 0.001$, $\eta^2_p = 0.139$, indicating that performance for this task was significantly better on the post-test. There was no significant moment by condition effect, $F(3,147) = 1.74$, $p = 0.161$, $\eta^2_p = 0.034$.

Also on the *connecting task*, a main effect of moment was present, $F(1,147) = 5.37$, $p = 0.002$, $\eta^2_p = 0.035$. There was a trend towards significant interaction between moment and condition, $F(3,147) = 2.38$, $p = 0.072$, $\eta^2_p = 0.046$ which was further explored with planned comparisons. These additional analyses revealed that the comparison condition improved significantly from pre- to post-test, $t(46) = 3.06$, $p = 0.004$, $d = 0.61$, which was not the case for the other conditions (NLE condition, $t(40) = 1.77$, $p = 0.085$, $d = 0.30$, active control condition, $t(36) = 0.057$, $p = 0.571$, $d = 0.09$, and empty control condition, $t(25) = 0.323$, $p = 0.749$, $d = 0.07$).

3.3.2. Comparison tasks

The performance on the *non-symbolic comparison task was improved on the post-test*, $F(1,147) = 6.51$, $p = 0.012$, $\eta^2_p = 0.042$. However, no significant differences between conditions were found as indicated by a non-significant interaction between moment and condition, $F(3,147) = 2.01$, $p = 0.115$, $\eta^2_p = 0.039$. A similar pattern was observed in the *symbolic comparison task*, where a main effect of moment was found, $F(1,147) = 6.65$, $p = 0.011$, $\eta^2_p = 0.043$, but no moment by condition effect, $F(3,147) = 1.72$, $p = 0.165$, $\eta^2_p = 0.034$.

Table 2
Correlations between all tasks.

	1	2	3	4	5	6	7
1. Counting task	–						
2. Connecting task	0.38**	–					
3. Non-symbolic comparison task	0.19*	−0.02	–				
4. Symbolic comparison task	0.23**	0.24**	0.20*	–			
5. Non-symbolic *NLE* task	−0.15	−0.10	−0.08	−0.10	–		
6. Symbolic *NLE* task	−0.22**	−0.11	−0.03	−0.26**	0.59**	–	
7. Arithmetic problems	0.25**	0.31**	0.11	0.44**	−0.08	−0.25**	–
8. Addition operations	0.31**	0.32**	0.11	0.38**	−0.28**	−0.28**	0.36**

*$p < 0.05$, **$p < 0.01$.

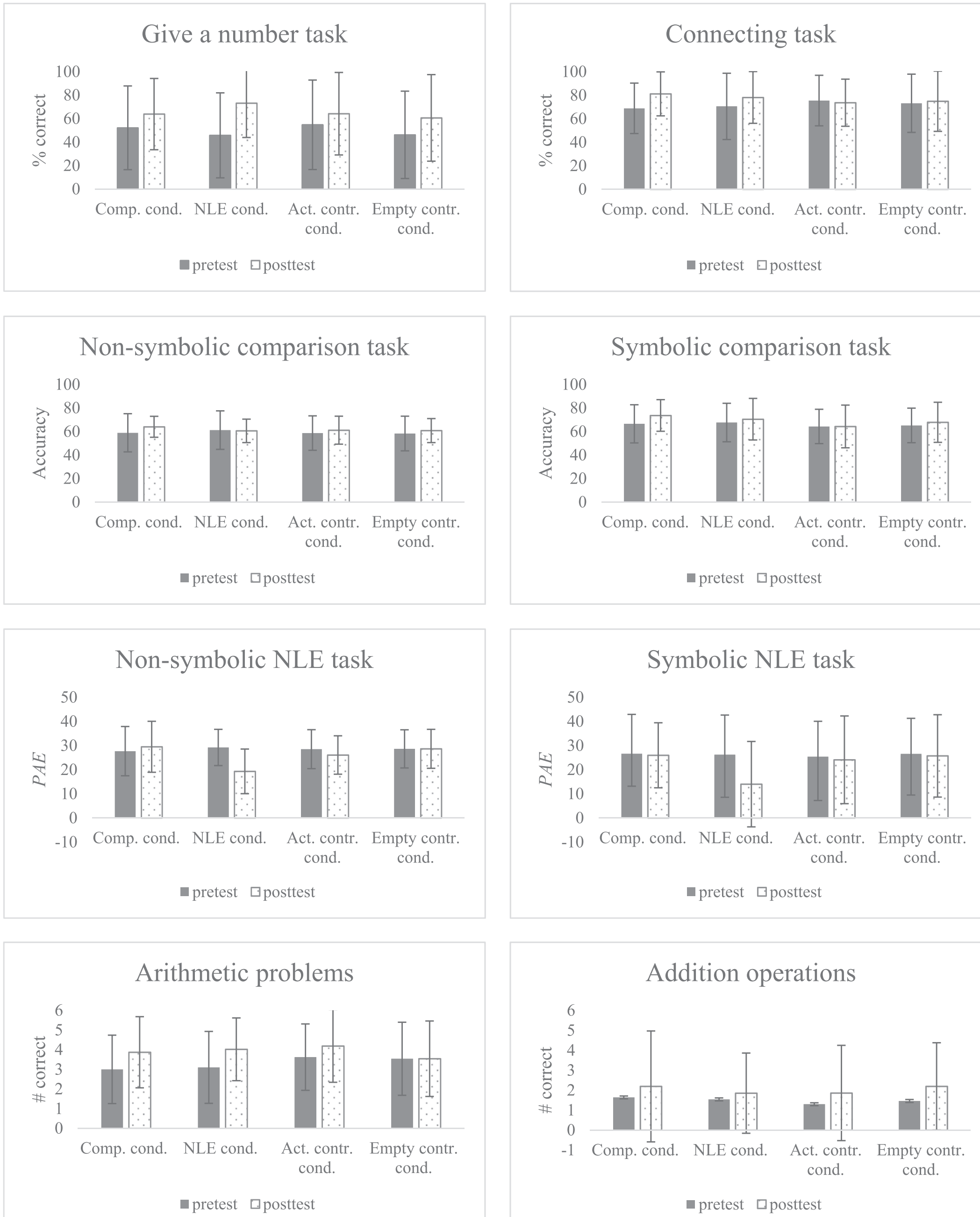

Fig. 5. Graphic representation of the mean scores (±SD) of all experimental tasks on pre- and post-test, per group.

3.3.3. Number line estimation tasks

On the *non-symbolic number line estimation task*, children's performance was better at the post-test than at the pre-test, $F(1,147) = 12.73$, $p < 0.001$, $\eta^2_p = 0.080$. An interaction effect between moment and condition was also observed, $F(3,147) = 14.22$, $p < 0.001$, $\eta^2_p = 0.225$. This interaction was further explored with pairwise comparisons of pre- and post-test. These revealed that the performance only improved significantly in the NLE condition, $t(40) = 6.72$, $p < 0.001$, $d = 1.17$ (comparison condition, $t(46) = 1.55$, $p = 0.129$, $d = 0.18$, active control condition, $t(36) = 1.63$, $p = 0.113$, $d = 0.30$, and empty control condition, $t(25) = -0.02$, $p = 0.984$, $d = 0.00$).

Similar results were obtained in the *symbolic number line estimation task*. A main effect of moment, $F(1,147) = 23.67$, $p < 0.001$, $\eta^2_p = 0.139$, indicating that performance was better in the post-test and a moment by condition interaction was observed, $F(3,147) = 14.74$, $p < 0.001$, $\eta^2_p = 0.231$. Here too, planned pairwise comparison of pre- and post-test revealed that only the NLE condition improved significantly from pre- to post-test, $t(40) = 6.93$, $p < 0.001$, $d = 1.15$ (comparison condition, $t(46) = 0.60$, $p = 0.551$, $d = 0.06$, active control condition, $t(36) = 0.92$, $p = 0.366$, $d = 0.14$, and empty control condition, $t(25) = 0.47$, $p = 0.641$, $d = 0.09$).

3.3.4. Arithmetic

On *arithmetic*, performance improved from pre to post-test resulting in a main effect of moment, $F(1,147) = 19.96$, $p < 0.001$, $\eta^2_p = 0.120$. There was a trend for a significant moment by condition interaction, $F(3,147) = 2.20$, $p = 0.090$, $\eta^2_p = 0.043$. Pairwise comparison of pre- and post-test showed that the performance in the comparison condition, $t(46) = -3.60$, $p = 0.001$, $d = 0.49$, and the NLE condition, $t(40) = -4.13$, $p < 0.001$, $d = 0.53$, improved significantly, which was not the case for the active, $t(36) = -2.46$, $p = 0.019$, $d = 0.32$, and empty control condition, $t(25) = 0.00$, $p = 1.00$, $d = 0.00$. For the *symbolic addition operations* a main effect of moment was observed, $F(3,147) = 17.54$, $p < 0.001$, $\eta^2_p = 0.107$, but no moment by condition interaction, $F(3,147) = 0.41$, $p = 0.745$, $\eta^2_p = 0.008$.

4. Discussion and conclusion

In this study, we contrasted two training programs that focused each on a specific basic number skill, i.e. comparison and number line estimation, in order to answer two outstanding questions in the domain of interventions on basic number processing, and, on the other hand, models of numerical cognition. First, we evaluated which training was more effective by examining their impact on arithmetic. Second, by contrasting both interventions, we gained more insight in the association between number comparison and number line estimation: do these skills rely on the same underlying representation as has been suggested or not?

When we compared the effect of both training programs on arithmetic performance, children improved in solving pictorially presented arithmetic problems, regardless of the numerical skill they trained. The performance of children improved significantly in both the number line estimation and in the comparison condition, with similar effects that were medium in their size. No significant improvements were observed in the active and empty control conditions. These results suggest that both comparison and number line estimation skills are causally related to arithmetic performance and confirm previous intervention studies in which comparison training (e.g., Obersteiner et al., 2013) or number line estimation training (e.g., Kucian et al., 2011; Ramani & Siegler, 2011) resulted in improved arithmetic. In the present study, we only observed a training effect for the pictorially presented arithmetic problems and not for the symbolic addition problems. In Flanders, formal instruction of symbolic addition problems is a part of the curriculum of the first year of elementary school and is typically not a part of the preparatory arithmetic in kindergarten (Vlaams Ministerie van Onderwijs en Vorming Agentschap voor Kwaliteitszorg in Onderwijs en Vorming, 2010). As a consequence, a straightforward explanation for this absence of an effect is that this task was too difficult for most of the children in our sample which was also evidenced by the low accuracies in that task.

We should be aware however, that the effects of number sense training on arithmetic were medium. It is widely recognized that arithmetic abilities are influenced by many factors (e.g. social, emotional and general cognitive factors). In this study, we focused on one of the factors that has been suggested as a building block for arithmetic is manipulated, i.e. number-specific skills. A recent meta-analysis (Schneider et al., 2016) showed that the association between symbolic magnitude comparison and arithmetic was only moderate (i.e., $r = 0.378$) and only explained about 14% of the variance in arithmetic. In light of this result, we cannot expect large effects when only number-specific skills are trained and interventions should probably focus on multiple factors. However, what this study clearly shows is that number-specific skills can certainly be one of these factors that is part of a large intervention as they lead to improved mathematical achievement.

Although both training programs resulted in higher arithmetic scores, they also resulted in different effects on the basic number skills that were measured. A first difference between both number sense trainings was the significantly increased performance on the connecting task after the comparison training, in contrast to no significant improvement after the number line estimation training. However, we believe that this observation has to be interpreted with caution. First, the effect size of improvement in the connecting task performance in the number line estimation condition was small ($d = 0.30$) indicating that also number line estimation training had some impact on the connecting task. This also means that the trend towards a moment by condition interaction was mainly driven by the absence of improvement in both control conditions. Second, this finding was not observed in a pilot study that contrasted the exact same trainings (Maertens, Sasanguie, De Smedt, Elen & Reynvoet, unpublished manuscript). As a result, we believe it is best to refrain from drawing strong conclusions based on this different impact of both number sense trainings on the connecting task.

More importantly, another difference between both number trainings was the observation that number line estimation skills improved after number line training (see also Link et al., 2013; Ramani & Siegler, 2008, 2011), but that these number line estimation skills were not boosted by practising comparison. This observation is at odds with the idea that both skills are relying on the same representation of number, i.e. the mental number line. The positioning of numbers on an external number line is considered as a reflection of the underlying mental number line (Booth & Siegler, 2006; Laski & Siegler, 2007). Changes in the number line estimation performance have been interpreted before as changes in the representation of numbers on the mental number line (e.g. a more linear representation, see Booth & Siegler, 2006). Also the performance in a comparison task and the developmental changes in comparison, have been explained in terms of changes in the number representations on the mental number line (e.g. Dehaene, 1997). Following this line of reasoning, improvement in performance of both tasks should go hand in hand. However, this was not the case in the present study, as we observed an asymmetry in the transfer effect. This asymmetry fits with Sasanguie and Reynvoet (2013), who found no relation between both tasks, and suggested that both tasks rely on different mechanisms. Recent studies suggested that a number line estimation task is based in part on spatial reasoning (Barth & Paladino, 2011; Link et al., 2013). For instance, children might use an anchor point strategy in which they initially use external anchor points (begin and end points) and later also use internal anchor points (e.g., middle of the number line) to position the numbers on a number line. These strategies are evidently not needed in the comparison task and may explain why no improvement is observed on the comparison tasks. We should note however, that this explanation remains largely speculative, as we did not directly examine strategy use in the number line estimation task. Nonetheless, the present results indicate that the improvements in the number line estimation task are probably not due to changes of the mental number line, because this would have led to improvements in comparison too.

A rather unexpected finding was that the performance on the comparison tasks in the post-test was statistically the same after comparison training compared to the other conditions. However, as can be inferred from Table 1, the gain scores on the comparison tasks were the highest in the comparison training condition, reflecting a similar pattern as the number line estimation performance. This was also reflected in effect sizes accompanying the pairwise comparison between pre and post-test performance we computed: the effect sizes (d) after comparison training were medium ($d = 0.50$ and $d = 0.47$ for non-symbolic and symbolic comparison tasks respectively) whereas they were small in the other conditions (all $ds < 0.25$).

A possible limitation of this study, that might perhaps be responsible for the lack of finding a significant effect on the comparison tasks even after comparison training, is the sample size. To verify whether the current sample size was large enough, we calculated post-hoc the power value of our study by means of the G*power program (Erdfelder, Faul, & Buchner, 1996) to detect intervention effects that have been reported in previous studies. For example, Siegler and Ramani (2009) reported that their linear board game had a medium effect ($\eta^2 = 0.19$). We calculated the power of our studies to detect a similar effect. Given the sample size in our study (i.e., four groups of 47, 41, 37, and 26 children) and the alpha level used ($\alpha = 0.05$), our study had a power of 0.99 to detect a medium effect. This analysis showed that the sample size in our study was most likely not the reason for the absence of effects.

Another explanation for the absence of effects in the comparison post-test may be the duration of the intervention. It remains possible that more extensive training would have led to larger improvements in comparison. At first sight this is at odds with a recent study of Hyde et al. (2014) in children that were on average about one year older. These authors demonstrated that during a brief training of only 60 trials, comparison performance improved already in the second half of the training. However, the improvement on comparison performance was only observed on reaction times and not on accuracies. The lack of effect on accuracies in the Hyde et al. (2014) study, was further evidenced by the observation that there were no differences between conditions in the numerical acuity, another accuracy related measure, that was measured after training. In sum, it remains possible that also our comparison training would have resulted in faster responding in the post-test comparison tasks. However these reaction times could not be analysed in the present study because performance was measured with tablets making the reaction times unreliable due to timing error (Schatz et al., 2015).

Finally, it is also possible that the lack of effect on comparison is due to reduced malleability (also in accuracy) of the comparison skills in contrast to the malleability of the number line estimation skills, leaving less room to improve on comparison.

To conclude, the present study with 5–6 year old kindergartners showed that the effectiveness of number line training and comparison training on arithmetic was very similar. A training of about 1 h spread over three weeks resulted in a (medium) positive effect on arithmetic. However, further training studies are needed to tackle the following outstanding issues. For instance, it needs to be determined whether a longer intervention results in larger effects and whether number sense trainings are equally effective in older children or in children with low SES. Currently, studies are lacking in order have an idea to what extend the effectiveness of training studies is influenced by the duration of the training and the age and/or background of the participants. In the future, it may also be interesting to investigate whether the effects of comparison and number line estimation training are additive. The absence of transfer effects from one task to another and the fact that both trainings influence arithmetic performance, suggests that they probably influence mathematics achievement through different mechanisms. A crucial test for this idea would be to combine both skills in one intervention and compare their effects with the current interventions focussing on one only skill.

References

Barth, H., La Mont, K., Lipton, J., & Spelke, E. S. (2005). Abstract number and arithmetic in preschool children. *Proceedings of the National Academy of Sciences, 102*(39), 14116–14121.

Barth, H., & Paladino, A. M. (2011). The development of numerical estimation: evidence against a representational shift. *Developmental Science, 14*, 125–135.

Berch, D. B. (2005). Making sense of number sense: Implications for children with mathematical disabilities. *Journal of Learning Disabilities, 38*, 333–339.

Berteletti, I., Lucangeli, D., Piazza, M., Dehaene, S., & Zorzi, M. (2010). Numerical estimation in preschoolers. *Developmental Psychology, 46*(2), 545–551.

Booth, J. L., & Siegler, R. S. (2006). Developmental and individual differences in pure numerical estimation. *Developmental Psychology, 41*, 189–201.

Borenstein. (2009). Effect sizes for continuous data. In H. Cooper, L. V. Hedges, & J. C. Valentine (Eds.), *The handbook of research synthesis and meta-analysis* (pp. 221–237). New York: Russell Sage Foundation.

Carey, S. (2009). *The origin of concepts*. Oxford: Oxford University.

Condry, K. F., & Spelke, E. S. (2008). The development of language and abstract concepts: The case of natural number. *Journal of Experimental Psychology: General, 137*(1), 22–38.

De Smedt, B., Noël, M. P., Gilmore, C., & Ansari, D. (2013). The relationship between symbolic and non-symbolic numerical magnitude processing and the typical and atypical development of mathematics. A review of evidence from brain and behaviour. *Trends in Neuroscience and Education, 2*, 48–55.

De Smedt, B., Verschaffel, L., & Ghesquière, P. (2009). The predictive value of numerical magnitude comparison for individual differences in mathematics achievement. *Journal of Experimental Child Psychology, 103*, 469–479.

Dehaene, S. (1997). *The number sense: How the mind creates mathematics*. London: The Penguim Press.

Dehaene, S., Izard, V., Spelke, E., & Pica, P. (2008). Log or linear? Distinct intuitions of the number scale in western and Amazonian indigene cultures. *Science, 320*(5880), 1217–1220.

Dieckman, N. F., Slovic, P., & Peters, E. M. (2009). The use of narrative evidence and explicit likelihood by decision makers varying in numeracy. *Risk Analysis, 10*, 1473–1488.

Erdfelder, E., Faul, F., & Bucher, A. (1996). GPOWER: a general power analysis program. *Behavior Research Methods, Instruments & Computers, 28*, 1–11.

Feigenson, L., Libertus, M. E., & Halberda, J. (2013). Links between the intuitive sense of number and formal mathematics ability. *Child Development Perspectives, 7*(2), 74–79.

Friso-van den Bos, I., van der Ven, S. H. G., Kroesbergen, E. H., & van Luit, J. E. H. (2013). Working memory and mathematics in primary school children: A meta-analysis. *Educational Research Review, 10*, 29–44.

Gallistel, C. R., & Gelman, R. (1992). Preverbal and verbal counting and computation. *Cognition, 44*, 43–74.

Gebuis, T., & Reynvoet, B. (2011). Generating nonsymbolic number stimuli. *Behavior Research Methods, 43*(4), 981–986.

GiggleUp Kids Apps And Educational Games Pty Ltd, 2013. Retrieved from: https://itunes.apple.com/be/app/mega-memory-match-het-grappige/id535014207?l=nl&mt=8.

Grégoire, J., Noël, M., & Van Nieuwenhoven, C. (2004). *TEDI-MATH*. Antwerpen: Harcourt.

Halberda, J., & Feigenson, L. (2008). Developmental change in the acuity of the "number sense": The approximate number system in 3-, 4-, 5-, and 6-year-olds and adults. *Developmental Psychology, 44*(5), 1457–1465.

Halberda, J., Mazzocco, M. M. M., & Feigenson, L. (2008). Individual differences in non-verbal number acuity correlate with maths achievement. *Nature, 455*, 1–11.

Holloway, I. D., & Ansari, D. (2009). Mapping numerical magnitudes onto symbols: The numerical distance effect and individual differences in children's mathematics achievement. *Journal of Experimental Child Psychology, 103*, 17–29.

Hunter, J. E., & Schmidt, F. L. (1990). *Methods of meta-analysis: Correcting error and bias in research findings*. Newbury Park, CA: Sage.

Hyde, D. C., Khanum, S., & Spelke, E. S. (2014). Brief non-symbolic, approximate number practise enhances subsequent exact symbolic arithmetic in children. *Cognition, 131*, 92–107.

Jordan, N. C., & Levine, S. C. (2009). Socioeconomic variation, number competence, and mathematics learning difficulties in young children. *Developmental Disabilities Research Reviews, 15*(1), 60–68.

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*(3), 782–795.

Laski, E. V., & Siegler, R. S. (2007). Is 27 a big number? Correlational and causal connections among numerical categorization, number line estimation, and numerical magnitude comparison. *Child Development, 78*(6), 1723–1743.

Lindskog, M., Winman, A., Juslin, P., & Poom, L. (2013). Measuring acuity of the approximate number system reliably and validly: The evaluation of an adaptive test procedure. *Frontiers in Psychology, 4*, 510. http://dx.doi.org/10.3389/fpsyg.2013.00510.

Link, T., Moeller, K., Huber, S., Fischer, U., & Nuerk, H. C. (2013). Walk the number line – An embodied training of numerical concepts. *Trends in Neuroscience and Education, 2*, 74–84.

Link, T., Nuerk, H. C., & Moeller, K. (2014). On the relation between the mental number line and arithmetic competencies. *The Quarterly Journal of Experimental Psychology, 67*(8), 1597–1613.

Linsen, S., Maertens, B., Husson, J., Van den Audenaeren, L., Wauters, J., Reynvoet, B., et al. (2015). Design of the game "Dudeman & Sidegirl: Operation clean world", a numerical magnitude processing training. In J. Torbeyns, E. Lehtinen, & J. Elen (Eds.), *Describing and studying domain-specific serious games* (pp. 9–26). Springer International Publishing.

Maertens, B., Sasanguie, S., De Smedt, B., Elen, J., & Reynvoet, B.. (unpublished manuscript). Comparison or number line estimation training: does it matter? Results of a number sense intervention with a tablet game.

Mundy, E., & Gilmore, C. K. (2009). Children's mapping between symbolic and nonsymbolic representations of number. *Journal of Experimental Child Psychology, 130*, 490–502.

Nieder, A., & Dehaene, S. (2009). Representations of number in the brain. *Annual Review of Neuroscience, 32*, 185–208.

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 Instruction, 23*, 125–135.

Park, J., & Brannon, E. M. (2014). Improving arithmetic performance with number sense training: An investigation of the underlying mechanism. *Cognition, 133*, 188–200.

Piazza, M., Facoetti, A., Trussardi, A. N., Berteletti, I., Conte, S., Lucangeli, D., et al. (2010). Developmental trajectory of number acuity reveals a severe impairment in developmental dyscalculia. *Cognition, 116*(1), 33–41.

Ramani, G. B., & Siegler, R. S. (2008). Promoting broad and stable improvements in low-income children's numerical knowledge through playing number board games. *Child Development, 79*(2), 375–394.

Ramani, G. B., & Siegler, R. S. (2011). Reducing the gap in numerical knowledge between low- and middle-income preschoolers. *Journal of Applied Developmental Psychology, 32*, 146–159.

Ramani, G. B., Siegler, R. S., & Hitti, A. (2012). Taking it to the classroom: Number board games as a small group learning activity. *Journal of Educational Psychology, 104*(3), 661–672.

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.

Reynvoet, B., Smets, K., & Sasanguie, D. (2016). Number sense: what's in a name and why should we bother? In A. Henik (Ed.), *Continuous issues in numerical cognition: How many or how much* (pp. 195–216). San Diego: Academic Press.

Sasanguie, D., De Smedt, B., Defever, E., & Reynvoet, B. (2012). Association between basic numerical abilities and mathematics achievement. *British Journal of Developmental Psychology, 30*, 344–357.

Sasanguie, D., Defever, E., Maertens, B., & Reynvoet, B. (2014). The approximate number system is not predictive for symbolic number processing in kindergartners. *The Quarterly Journal of Experimental Psychology, 67*(2), 271–280.

Sasanguie, D., Göbel, S., Moll, K., Smets, K., & Reynvoet, B. (2013). Acuity of the approximate number sense, symbolic number comparison or mapping numbers onto space: What underlies mathematics achievement? *Journal of Experimental Child Psychology, 114*(3), 418–431.

Sasanguie, D., & Reynvoet, B. (2013). Number comparison and number line estimation rely on different mechanisms. *Psychologica Belgica, 53*(4), 17–35.

Schatz, P., Ybarra, V., & Leitner, D. (2015). Validating the accuracy of reaction time assessment on computer-based tablet devices. *Assessment, 22*, 405–410.

Schneider, M., Beeres, K., Coban, L., Merz, S., Schmidt, S. S., Stricker, J., et al. (2016). Associations of non-symbolic and symbolic numerical magnitude processing with mathematical competence: A meta-analysis. *Developmental Science*. http://dx.doi.org/10.1111/desc.12372 (Epub ahead of print).

Siegler, R. S., & Booth, J. L. (2004). Development of numerical estimation in young children. *Child Development, 75*(2), 428–444.

Siegler, R. S., & Ramani, G. B. (2009). Playing linear number board games – But not circular ones – Improves low-income preschoolers' numerical understanding. *Journal of Educational Psychology, 101*(3), 545–560.

Vanbinst, K., Ghesquière, P., & De Smedt, B. (2015). Does numerical processing uniquely predict first graders' future development of single-digit arithmetic? *Learning & Individual Differences, 37*, 153–160.

Vlaams Ministerie van Onderwijs en Vorming, Agentschap voor Kwaliteitszorg in Onderwijs en Vorming. (2010). *Ontwikkelingsdoelen en eindtermen voor het gewoon basisonderwijs. Informatie voor de onderwijspraktijk* [*Developmental objectives and attainment targets for regular elementary education. Information for educational practice*]. Retrieved from: http://www.vlaanderen.be/nl/publicaties/detail/ontwikkelingsdoelen-en-eindtermen-voor-het-gewoon-basisonderwijs-informatie-voor-de-onderwijspraktijk.

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*(4), 224–234.

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), 1–14.