Journal of Experimental Child Psychology 116 (2013) 775–791

Contents lists available at ScienceDirect

Journal of Experimental Child Psychology

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

Developing number–space associations: SNARC effects using a color discrimination task in 5-year-olds ☆

Danielle Hoffmann *, Caroline Hornung, Romain Martin, Christine Schiltz

Educational Measurement and Applied Cognitive Science (EMACS) Unit, Faculty of Language and Literature, Humanities, Arts, and Education (FLSHASE), University of Luxembourg, L-7201 Walferdange, Luxembourg

A R T I C L E I N F O

Article history:
Received 4 June 2013
Revised 22 July 2013

Keywords:
SNARC
Numerical cognition
Number–space associations
Development
Symbolic numbers
Magnitude representation

A B S T R A C T

Human adults' numerical representation is spatially oriented; consequently, participants are faster to respond to small/large numerals with their left/right hand, respectively, when doing a binary classification judgment on numbers, known as the SNARC (spatial–numerical association of response codes) effect. Studies on the emergence and development of the SNARC effect remain scarce. The current study introduces an innovative new paradigm based on a simple color judgment of Arabic digits. Using this task, we found a SNARC effect in children as young as 5.5 years. In contrast, when preschool children needed to perform a magnitude judgment task necessitating exact number knowledge, the SNARC effect started to emerge only at 5.8 years. Moreover, the emergence of a magnitude SNARC but not a color SNARC was linked to proficiency with Arabic digits. Our results suggest that access to a spatially oriented approximate magnitude representation from symbolic digits emerges early in ontogenetic development. Exact magnitude judgments, on the other hand, rely on experience with Arabic digits and, thus, necessitate formal or informal schooling to give access to a spatially oriented numerical representation.

* Corresponding author. Fax: +352 46 66 44 9453.
E-mail address: danielle.hoffmann@uni.lu (D. Hoffmann).

0022-0965/$ - see front matter © 2013 The Authors. Published by Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.jecp.2013.07.013

Introduction

Modern society expects us to acquire a complex and sophisticated understanding of numbers, quantities, and their manipulations. To foster this cognitive ability and develop adequate mathematics curricula, it is critical to understand how basic number skills arise and mature in typically developing children. Understanding how number–space associations develop is one of these key challenges (de Hevia, Girelli, & Macchi Cassia, 2012).

Healthy adults in Western cultures are known to associate small numbers with the left side of space and large magnitudes with the right side of space (for reviews, see de Hevia, Vallar, & Girelli, 2008; Fias & Fischer, 2005). The most well-known evidence for the spatial quality of numerical information is the so-called SNARC (spatial–numerical association of response codes) effect first described by Dehaene and colleagues during the early 1990s (Dehaene, Bossini, & Giraux, 1993; Dehaene, Dupoux, & Mehler, 1990). This effect refers to the finding that participants tend to respond faster with their left hand to small numbers and faster with their right hand to large numbers when doing binary classification tasks on single digits (Dehaene et al., 1990, 1993). It is often interpreted as revealing a spatial code in which numbers are represented horizontally (Restle, 1970) and (at least in Western participants) from the left to the right (Dehaene et al., 1993) according to their magnitude. Recently, it has also been shown that verbal coding (Gevers et al., 2010; Proctor & Cho, 2006) and processing of ordinal information in working memory (van Dijck & Fias, 2011; see also Previtali, de Hevia, & Girelli, 2010) play important roles in the SNARC effect. The association of small/large numbers with the left/right side of space, respectively, can be found using classification tasks requiring an explicit access of number magnitude such as a magnitude judgment task (Dehaene et al., 1990). Intriguingly, the same effect is also observed when the classification task does not require explicit magnitude access, for instance, when using a parity judgment task (Dehaene et al., 1993). But even though the SNARC effect has been reliably tested in adults (Wood, Willmes, Nuerk, & Fischer, 2008), surprisingly few developmental studies exist.

The age when number symbol processing starts revealing spatial hallmarks was first investigated in a study by Berch, Foley, Hill, and Ryan (1999). These authors found a SNARC effect only in children aged 9 years and onward, but not in 8-year-olds, using a parity judgment task (see also Schweiter, Weinhold Zulauf, & von Aster, 2005). A recent study by White, Szűcs, and Soltész (2011) extended the findings of the study by Berch and colleagues (1999) by increasing experimental power. Using a parity judgment task adapted to young children on parameters such as feedback and stimuli used, the authors were able to detect a SNARC effect already in their 7.5-year-old participants but not in their 6.2-year-old participants. Another study by van Galen and Reitsma (2008) investigated the SNARC effect in young children using both a magnitude judgment task and a lateral target detection task following nonpredictive digit cues. In this task, a single digit is used as a central nonpredictive cue, and the general finding is that participants are faster to detect a left-sided target after a small digit and a right-sided target after a large digit (Fischer, Castel, Dodd, & Pratt, 2003). Van Galen and Reitsma found a SNARC effect in the magnitude judgment task in all three of their child groups (7-, 8-, and 9-year-olds), whereas in the detection task they found an effect in the 9-year-olds but not in the younger children. Taken together, these results show that children associate small and left as well as large and right already at 7 years of age when tested in tasks requiring digit processing.

The failure to find SNARC effects in children younger than 7 years is surprising given that preliterate children already order magnitude information from the left to the right, at least in Western cultures (Opfer, Thompson, & Furlong, 2010). In their first experiment, Opfer and colleagues (2010) presented a group of 4-year-old children with a spatial search task. The task involved a box with seven compartments labeled with the numerals 1 to 7. For half of the children the numbers increased from left to right, and for the other half the numbers increased from right to left. The results showed that children were faster and more accurate in the left-to-right search condition; they appeared to have a robust expectation for numbers to increase in a left-to-right manner. In their second experiment, the authors asked preschool children (3–5 years of age) to count a given linear set of objects. Here they found that already 73% of the preschoolers started counting at the leftmost object and proceeded to the rightmost object. Subsequently, the children were asked to add an object to the set and later to subtract an object from the set. Approximately 40% of the 5-year-olds both added from the left to the right (by placing an additional object to the right of the set) and subtracted from the right to the left (by taking away one of the objects of the set from the right side).

A recent study by Shaki, Fischer, and Göbel (2012) confirmed and extended these findings by showing that already 4-year-old preschoolers counted in the reading direction that was habitual in their culture. Whereas 60.7% of tested British preschoolers counted from the left to the right, 66.2% of the Palestinian preschoolers counted from the right to the left. The directional bias increased with formal schooling, with 91.3% of the British and 78.7% of the Palestinian 9-year-olds counting in their respective cultural reading direction. Because illiterate adults showed no directional preference, the authors concluded that a very likely candidate source for the culturally determined direction in spatio–numerical associations is observational learning. Even before formally learning to read, children are able to observe from a very young age the habitual scanning strategies in use in a given culture when living in a literate society. Finally, another important hint for the early presence of number–space interactions comes from an elegant study using a nonsymbolic SNARC paradigm in preschoolers (Patro & Haman, 2012). In this experiment, 4-year-olds were indeed systematically faster when choosing the display containing less/more elements when it was presented on the left/right side of the screen, respectively.

Preschool is a critical phase for numerical development in general and for number symbol mastery in particular. Although already infants (and even nonhuman animals) possess core systems for large approximate and small exact number representation (Feigenson, Dehaene, & Spelke, 2004; Noël & Rousselle, 2011; Siegler & Opfer, 2003; Xu & Spelke, 2000), these systems are limited in their representational scope. To allow exact processing of large numerosities above three and achieve the mature numerical performance of educated human adults, these core systems need to be extended by the acquisition of exact number words and symbols. For connecting these newly acquired number symbols to the core systems, preschool is a stage of particular interest (Booth & Siegler, 2006; Lipton & Spelke, 2005; Siegler & Booth, 2004). Thus, several studies from independent groups reported that preschoolers' exact knowledge of number symbols is systematically related to their approximate number system's (ANS's) acuity (Gilmore, McCarthy, & Spelke, 2010; Libertus, Feigenson, & Halberda, 2011; Mussolin, Nys, Leybaert, & Content, 2012; but see also Piazza et al., 2010; Sasanguie, Göbel, Moll, Smets, & Reynvoet, 2013; Soltész, Szucs, & Szucs, 2010). Moreover, older preschool children (attending the last grade of kindergarten) are also able to use (even two-digit) number symbols for approximate numerical computations such as estimation of large item sets (Huntley-Fenner, 2001; Mejias & Schiltz, 2013) or approximate symbolic calculations (Gilmore, McCarthy, & Spelke, 2007).

When we combine the observation that preschoolers can access approximate magnitude representations from number symbols with the fact that they show left-to-right preferences when counting and processing numerosities, it follows that SNARC effects associated with number symbols should already appear in preschool children if tested appropriately. Especially tasks that do not need exact numerical computation should lead to a SNARC effect in preschoolers because they activate approximate number representations that are thought to be spatially oriented (e.g., de Hevia et al., 2012).

To address this hypothesis, the current study assessed SNARC effects in preschool children using both a magnitude-irrelevant task and a magnitude-relevant task on Arabic digits (see also van Galen & Reitsma, 2008). The magnitude-irrelevant task gives indications about the automatic processing of the magnitude information and, hence, provides a picture of the internal representation that is uncontaminated by intentional operations (Tzelgov & Ganor-Stern, 2005). This aspect is particularly important when studying numerical development in preschoolers because they are just elaborating the abstract meaning associated with number symbols, and explicit access to these representations is still effortful and requires the investment of important cognitive resources. Therefore, at this developmental stage, the implicit magnitude task, more than the explicit magnitude task, is expected to reveal spatio–numerical interactions associated with the automatic activations of the spatially oriented ANS representations. Previous SNARC studies used parity judgment tasks to investigate the development of number–space associations, but successfully judging the parity status of numbers requires formal math schooling and, therefore, might be suboptimal for testing young children at the beginning of primary school and earlier (Berch et al., 1999; Fias & Fischer, 2005; van Galen & Reitsma, 2008; White, Szűcs, & Soltész, 2011). To avoid this methodological limitation, we developed a new magnitude-irrelevant SNARC paradigm requiring an easy binary classification task. Children simply needed to discriminate whether an Arabic digit was presented in red or green ink. To allocate sufficient time for automatic semantic processing of the digit magnitude, the red/green color appeared after an initial 200-ms period of black digit presentation. Our task was methodologically similar to other magnitude implicit tasks such as the phoneme detection task (Fias, Brysbaert, Geypens, & d'Ydewalle, 1996) and the vowel/consonant judgment or symmetry/asymmetry judgment tasks (Huha, Berch, & Krikorian, 1995, cited in Fias & Fischer, 2005) previously employed in adults, but with a classification judgment adapted to very young children.

In addition, we used a classical magnitude judgment task (e.g., van Galen & Reitsma, 2008) to assess spatio–numerical interactions during exact processing of number symbols. This task, moreover, allowed computing a distance effect. Distance effects denote the finding that participants are slower when comparing two numbers that are close to each other than when comparing two numbers that are further apart (Moyer & Landauer, 1967), and they have been described in children as young as 5 years (Duncan & McFarland, 1980; Sekuler & Mierkiewicz, 1977). To further evaluate numerical competencies of the participating preschool children and see whether they influence spatio–numerical representations, we also tested counting and digit writing abilities and administered a number estimation task. In the number estimation task, we adopted the logarithmic-to-linear shift framework (Opfer, Siegler, & Young, 2011; Siegler & Opfer, 2003). According to this account, a qualitative (from logarithmic to linear) shift of numerical representations occurs over development (Opfer et al., 2011; Siegler & Opfer, 2003). Alternatively, segmented linear models propose distinct linear representations for small and large number ranges based on familiarity with number ranges (Ebersbach, Luwel, Frick, Onghena, & Verschaffel, 2008) or mastery of the Arabic place-value system (Moeller, Pixner, Kaufmann, & Nuerk, 2009). Recently, power models have also been proposed to describe number estimation data, attributing observed developmental changes in number estimation to a reduction of estimation bias with age and experience (Barth & Paladino, 2011; Slusser, Santiago, & Barth, 2013).

In summary, there is ample evidence for number–space associations assessed as SNARC effects in young healthy adults; however, knowledge concerning their developmental trajectory is much sparser. Especially the question of when spatial effects related to digit processing first arise needs further investigation. In the current study, we tested preschool children to evaluate number–space associations before the start of formal literacy and numeracy education. We used a newly developed digit color judgment paradigm as well as a classical magnitude judgment task. To closely monitor the development of number–space associations during this crucial preschool period (Booth & Siegler, 2006; Lipton & Spelke, 2005; Siegler & Booth, 2004), two groups of children were included in the study. The first-term group was tested at the beginning of the last year of kindergarten and was composed of 5.5 (± 0.3)-year-old children, whereas the second-term group was composed of 5.8 (± 0.3)-year-old children who were tested in the middle of the year. In line with reports of early predispositions to associate numerosity and space (de Hevia & Spelke, 2010; Patro & Haman, 2012), we hypothesized that both groups of preschool children would reveal significant SNARC effects in the digit color judgment task. Because it does not require explicit magnitude judgments, it should indeed highlight the effects of an automatic activation of spatially oriented magnitude representations. In the explicit magnitude judgment task, in contrast, we expected that the children's mastery of number symbols would heavily influence the strength of their SNARC effects. The more preschoolers rely on mature number processing in this task requiring access to exact number concepts, the stronger their expected SNARC effects.

Method

Participants

A total of 84 children from the last year of kindergarten in Luxembourgish schools participated in the current study. One group was tested at the beginning of the school year (in the first term, $n = 36$, mean age = 5.53 ± 0.31 years, 18 girls and 18 boys, 1 left-handed), and the second group was tested in the middle of the school year (in the second term, $n = 48$, mean age = 5.84 ± 0.32 years, 24 girls and 24 boys, 3 left-handed).

Materials and procedure

The computerized tasks were programmed in E-Prime (Version 2.0.8.79; Schneider, Eschmann, & Zuccolotto, 2002) and administered using a Lenovo ThinkPad 61 Tablet Laptop with a 12.1-inch color monitor (1024 × 768 pixels). A paper mask with two holes for the answer keys was used during task administration in order to avoid distraction by other keys and to cover up the number line of the keyboard.

Experimental tasks

Magnitude judgment task: Explicit magnitude access. During the magnitude judgment task, the children needed to judge whether a centrally presented Arabic digit was smaller or larger than 5. This task allows testing for the presence of a SNARC effect and constitutes a task with explicit magnitude information access because exact digit magnitude is central for the correct completion of the task. Furthermore, it allows detecting the presence of a distance effect.

The design of this task was adapted from van Galen and Reitsma (2008). Each trial started with an empty black-bordered square on a white background (sides 100 pixels, border 2 pixels). After 1000 ms, one of eight possible stimuli (Arabic digits 1, 2, 3, 4, 6, 7, 8, and 9), presented in black on a white background in font Arial point size 48, appeared at the center of the square and remained until response or until 5000 ms had elapsed. The intertrial interval was a blank screen of 1000 ms. The stimuli were presented in a pseudo-random order, no number appeared twice in a row, and the correct response could be on the same side no more than two times consecutively. Responses were given by pressing the "A" or "L" key of a standard QWERTZ keyboard, using the forefinger of the left hand or right hand, respectively. A cartoon mouse or elephant tag was placed above the answer key to indicate "smaller" or "larger," respectively.

Each child completed two blocks, one in each mapping (in one block "A" was assigned to "smaller than 5" and associated with the mouse tag, and in the other one "A" was assigned to "larger than 5" and associated with the elephant tag); block order was counterbalanced across participants. Each block started with 8 to 16 training trials, depending on response accuracy. An accuracy threshold of 80% correct answers needed to be reached in order to proceed directly to the experimental trials after 8 training trials; if the threshold was not reached, another 8 training trials were administered before the experimental trials started. The experiment itself consisted of 80 trials, 40 trials per block; mapping was changed during the break between the blocks.

Digit color judgment task: Implicit magnitude access. The parameters of the digit color judgment task were similar to those of the magnitude judgment task. During the color judgment task, the children needed to decide whether a centrally presented Arabic digit was red or green.

The central digit was first presented in black for 200 ms and then switched to either red or green. For half of the children the key assignment was to press "A" when the digit switched to red and "L" when the digit switched to green, and vice versa for the other half of the children. The experiment consisted of 80 experimental trials preceded by 8 to 16 training trials, depending on response accuracy. As before, a threshold of 80% correct responses needed to be reached to proceed to the experimental trials after only 8 training trials. The mouse and elephant tags were replaced by red- and green-colored tags, respectively. All other task parameters were identical to those in the magnitude judgment task. In the color judgment task, digit magnitude is irrelevant for successful task completion; hence, the presence of a SNARC effect in this task points toward an automatic activation of magnitude information.

General numerical assessment

Verbal counting task. The verbal counting ability of each child was tested using four subtests from the Diagnostic Tests of Metacognitions and Mathematics (Salonen et al., 1994, cited in Aunola, Leskinen, Lerkkanen, & Nurmi, 2004).

In the first subtest, the task consisted of counting as far as possible. If the child reached 50, the subtest was stopped. The child was given 1 point for counting correctly to 10, 2 points for reaching 20 without errors, and 3 points for reaching 50 without errors.

In the second subtest, a starting number was given from which the child was asked to count upward. Starting numbers were 3, 8, 12, and 19. The child was given 1 point if he or she counted upward correctly for at least four numbers from the starting number.

The third subtest consisted of counting backward from a given starting number. Starting numbers were 4, 6, 12, and 23. The child was given 1 point if he or she counted backward correctly for at least four numbers from the starting number.

In the fourth subtest, the child was asked four questions related to knowledge of the ordinal chain of numbers (e.g., "What number do you get when you count three numbers up from 5?"). The child was given 1 point for each correct response.

A sum score per child was computed by adding the scores of the four subtests, allowing a possible maximum of 15 points. In addition, we were interested in the results of the first subtest, counting to a maximum, for which the score corresponded to the last numeral counted to without any mistakes or any misses. This was recorded separately to have a comparable verbal counterpart to the written counting task described below.

Arabic digit writing task. The child was asked to write down the Arabic numerals as far as he or she knew on a blank white sheet of DinA4 paper. The child was given 1 point for each correct number; if one numeral was missing, the score corresponded to the last correct digit in the sequence.

Number line task (number-to-position task). The number line task was adapted from the number-to-position task described by Siegler and Opfer (2003). The children were presented with six different problems in the number range from 0 to 20, each presented on a separate sheet. The number line task allows appreciating the logarithmic versus linear nature of children's representation of the evaluated magnitude range. A logarithmic representation would reflect a more immature representation than a linear one because, according to the logarithmic-to-linear shift account, a qualitative (from logarithmic to linear) shift in numerical representations occurs over development (Siegler & Opfer, 2003; but see Barth & Paladino, 2011, for an alternative account).

Each sheet contained one black 25-cm-long horizontal line, with the left end labeled "0" and the right end labeled "20". Each number to be estimated (2, 4, 6, 12, 15, or 18, always in that order) was presented 2 cm above the center of the line. The child needed to indicate by putting a pencil mark on the line where the target number should be positioned between 0 and 20.

Magnitude judgment task: Distance effect. Using the magnitude judgment task, which is a comparison judgment task between two numerals, the distance effect can be computed.

The children were tested individually in a separate and quiet room in their school. To ensure that the children would be able to do the experimental tasks correctly, each child was asked verbally whether a given number (8, 6, 3, 1, 7, 4, 9, or 2, always in that order) was smaller or larger than 5 before starting the experimental tasks. Correct feedback was given, and the task was repeated until fewer than two mistakes were made.

After this initial training, half of the children started with the magnitude judgment task followed by the color judgment task, and vice versa for the other half of the children. The computerized tasks were followed by the verbal counting task, the Arabic digit writing task, and the number line task, always in that order. Testing took approximately 25 min per child.

Statistical analysis

Prior to data analysis, incorrect trials (with respect to magnitude or color judgment) and outliers were removed from the data. A trial was considered an outlier if the reaction time (RT) diverged from the participant's individual mean by 2.5 standard deviations. All data of a participant were excluded from group analysis if the child failed to reach an accuracy level of 75% correct answers.

Following these guidelines, 6.98% of all trials in the magnitude judgment task were removed due to errors and 2.54% were removed because they qualified as outliers. In the color judgment task, 4.68% of all trials were error trials and 2.93% were outliers and consequently removed. In the magnitude judgment task 14 of 84 children failed to reach an accuracy level of 75% correct answers, whereas in the color judgment task 3 of 84 children failed to reach the predefined accuracy threshold. The trimmed data of the 70 remaining participants in the magnitude judgment task and the 81 remaining participants in the color judgment task were subsequently submitted to statistical analysis.

To test for the presence of SNARC effects, we conducted a $2 \times 2 \times 2 \times 2$ repeated-measures analysis of variance (ANOVA) on RTs for each task separately, including response side (left or right) and digit magnitude (small [1–4] or large [6–9]) as within-participants factors and group (first term or second term) and task order (magnitude–color judgment or color–magnitude judgment) as between-participants factors. A SNARC effect would reveal itself as an interaction between response side and digit magnitude.

In addition, we used linear regression analysis methods for repeated-measures data following Lorch and Myers (1990), as suggested by Fias and colleagues (1996). For each participant, we computed mean RTs for each digit and each response side separately and then calculated individual difference scores (dRTs) by subtracting left-handed RTs from right-handed RTs for each digit. These dRTs were submitted to a linear regression analysis, using digit magnitude as a predictor variable.

This method allows obtaining for each participant a regression weight reflecting the size and direction of number–space associations. A significantly negative regression weight indicates the presence of a SNARC effect, meaning a number–space association in the expected direction (small numbers associated with the left side and large numbers associated with the right side). Among other advantages (for a detailed discussion, see Fias, Brysbaert, Geypens, & d'Ydewalle, 1996), this method has the benefit of providing a separate score (regression weight) for each participant, allowing us to recognize the presence or absence of a SNARC effect individually. In contrast, by using an ANOVA, the presence of the SNARC effect must be obtained from an interaction between two factors (response side and digit magnitude). Having an individual score allows correlating the individual SNARC effect with other individual measures such as counting ability.

The distance effect was computed using a linear regression analysis. Mean RT was computed for each child and each digit separately and then was regressed using the distance to the reference "5" as the predictor variable. The resulting regression weights were tested against "0" using a one-sided t test. A significantly negative slope indicates the presence of a distance effect.

Results and discussion

General numerical assessment

Details of the descriptive information of the participants as well as the scores of the general assessment tasks can be found in Table 1.

The two preschool children groups did not differ in gender or handedness distribution, $\chi^2(1) = 0$, $p = 1.00$, or $\chi^2(1) = 0.55$, $p > .60$, respectively. As expected, they differed significantly in mean age, with the children tested in the first term being 3.74 months younger than the children tested in the second term, $t(82) = -4.42$, $p < .001$.

In the general numerical assessment, the groups neither differed in the verbal counting task including all subtests, $t(81) = -0.02$, $p > .90$, nor differed when considering only the first subtest where they needed to count as far as they could, $t(81) = -0.34$, $p > .70$. However, they differed significantly in the Arabic digit writing task; on average, the children tested in the second term knew how to write Arabic digits nearly up to 7, whereas the children tested in the first term already stopped at around 4, $t(82) = -2.16$, $p < .05$. In the number line task, the younger children of the first-term group displayed a significantly better logarithmic fit than linear fit to account for their number estimations, $t(5) = 2.61$, $p < .05$. For the older children of the second-term group, the estimation data could just as well be approximated with a linear model as with a logarithmic model, $t(5) = 0.55$, $p > .60$. The two groups did not differ between each other on the linear fit, $t(10) = 1.82$, $p = .10$, or on the logarithmic fit, $t(10) = 0.60$, $p > .50$.

In the magnitude judgment task, the children tested in the first term were on average a significant 62.74 ms faster to respond to small numbers than to large numbers, $t(28) = 2.80$, $p < .01$, whereas the

Table 1
Descriptive information and mean performance for the two groups in the general numerical assessment tasks.

	First-term group [mean (*SD*)]	Second-term group [mean (*SD*)]	*t*
Descriptive information			
N	36	48	
Gender (male/female)	18/18	24/24	$\chi^2(1) = 0.00$
Age (months)	66.36 (3.77)	70.1 (3.89)	−4.42**
Handedness (right/left)	35/1	45/3	$\chi^2(1) = 0.55$
Verbal counting task			
Counting to maximum	16.31/50 (12.64)	17.27/50 (12.65)	−0.34
Total	5.89/15 (2.55)	5.90/15 (3.27)	−0.02
Arabic digit writing	4.44 (4.67)	6.67 (4.66)	−2.16*
Number-Line Task			
Linear fit (R^2)	.90† ($p < .05$)	.97	1.82
Logarithmic fit (R^2)	.97† ($p < .05$)	.98	0.61
Magnitude judgment task			
Distance effect	−40.04 (52.03)** from 0	−45.51 (52.55)** from 0	0.4

Note. Standard deviations are shown in parentheses. RTs are given in milliseconds (ms). For the number line task, R^2 values of the linear and logarithmic models fit on the data of the entire respective group are given, as suggested by Opfer's (2003) tutorial for this task. A significant difference between the linear and logarithmic fits within one group is indicated by a dagger (†) followed by the significance level.
* $p < .05$.
** $p < .001$.

older children were only 2 ms faster to respond to small numbers than to large numbers, $t(40) = 0.24$, $p > .80$. This finding corresponds to a size effect present in the first-term children only. A size effect (Siegler & Robinson, 1982) consists of faster RTs when comparing small numbers (e.g., 2 vs. 3) than when comparing large numbers (e.g., 8 vs. 9) that are separated by an equal distance. To better appreciate this size effect, we computed a linear regression analysis on RTs with digit magnitude as a predictor variable. This analysis revealed that in the first-term group only, RTs increased linearly 11.8 ms per digit; a two-sided *t* test confirmed that the slope differed significantly from 0, $t(28) = 2.89$, $p < .01$. This linear increase in RTs suggests that children in the first-term group used a serial counting strategy in order to be able to fulfill the task requirements. Similar to a serial scanning strategy, a counting

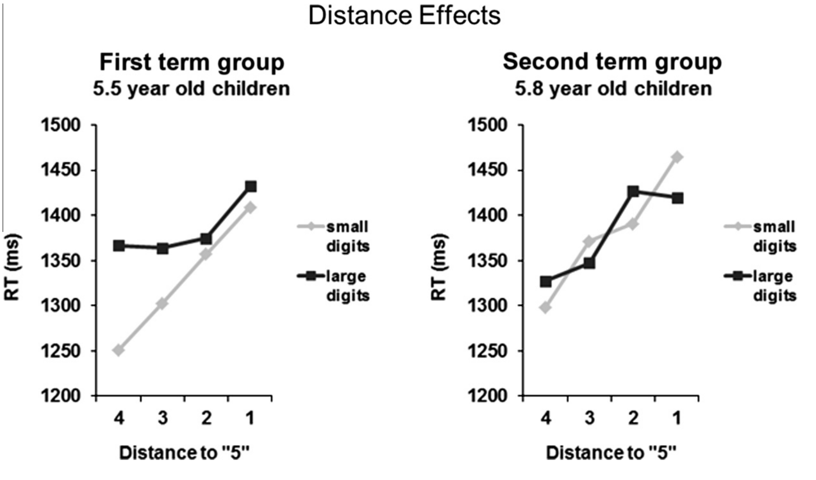

Fig. 1. Mean RTs for each digit separately for the two groups, illustrating distance effects in both groups.

strategy would involve linearly increasing RTs for each object in the series, leading to relatively shorter RTs when deciding that 2 comes before 5 in the counting routine than when deciding that 7 comes after 5. In the second-term group, RTs stayed constant over the range of digits, $t(40) = 0.46$, $p > .60$ (see also Carey, 2001, and van Dijck & Fias, 2011, for similar interpretations of monotonically increasing RTs with each n in a series).

A linear regression analysis on RTs for each digit using the distance to the reference (5) as a predictor variable allowed testing for the presence of a distance effect. This analysis revealed a significantly negative regression weight (–43.24, $p < .001$ compared with 0, two-tailed t test), demonstrating the presence of a distance effect (see Fig. 1). A one-way ANOVA revealed that the distance effect did not significantly differ between the first- and second-term groups, $F(1,68) = 0.19$, $p > .60$.

In sum, these measures indicate that numerical processing of the first-term children was less mature than the second-term children's performance. Indeed, the presence of a significant size effect indicated that the first-term group used serial counting procedures to solve the magnitude judgment task. Moreover, they could write fewer digits and relied more on logarithmic representations than on linear representations when placing numbers on a number line ranging from 0 to 20.

Number–space associations: Magnitude judgment task

Factorial analysis (ANOVA)

In the magnitude judgment task, the four-way mixed ANOVA with response side (left or right) and digit magnitude (small [1–4] or large [6–9]) as within-participants factors and group (first term or second term) and task order (magnitude–color judgment or color–magnitude judgment) as between-participants factors revealed a main effect of digit magnitude, $F(1,66) = 5.36$, $p < .05$, $\eta^2 = .075$. The magnitude factor interacted with group, $F(1,66) = 3.96$, $p = .05$, $\eta^2 = .06$, reflecting the size effect present in the first-term children described before. No other effects reached significance (all $ps > .05$). Most important for the current study, there was no interaction between response side and digit magnitude, $F(1,66) = 0.15$, $p > .70$, $\eta^2 = .002$, and, hence, no significant SNARC effect.

To capture developmental differences, we also performed a two-way ANOVA with the factors response side (left or right) and digit magnitude (small or large) as within-participants factors for each group separately.

In the first-term group, only the main effect of magnitude was significant, $F(1,28) = 6.50$, $p < .05$, $\eta^2 = .19$. No other effects reached significance; most important, the interaction between response side and digit magnitude did not reach significance either, $F(1,28) = 0.93$, $p > .30$, $\eta^2 = .03$. In the second-term group, no effects reached statistical significance, but the interaction between response side and digit magnitude tended toward significance, $F(1,40) = 2.90$, $p < .10$, $\eta^2 = .07$. In other words, there was no SNARC effect in the younger children, but there was a trend toward a SNARC effect in the older children.

Linear regression analysis

The linear regression analysis on dRTs (right-handed RT – left-handed RT) yielded a regression weight of –5.50 for the entire population in the magnitude judgment task. A one-tailed t test confirmed the absence of significant SNARC in the magnitude judgment task because the regression weight of the dRT slope was not significantly negative, $t(69) = -0.50$, $p > .30$ (see Fig. 2). A univariate ANOVA on regression weights, including group and task order as between-participants factors, revealed no effects reaching significance.

Number–space associations: Digit color judgment task

Factorial analysis (ANOVA)

In the digit color judgment task, the four-way mixed model ANOVA also revealed a significant effect of digit magnitude, $F(1,77) = 4.16$, $p < .05$, $\eta^2 = .05$, due to the children being faster to respond to large numbers (6–9) than to small numbers (1–4) (RTs = 1106.7 and 1085.6 ms for small and large numbers, respectively). No other main effects reached significance.

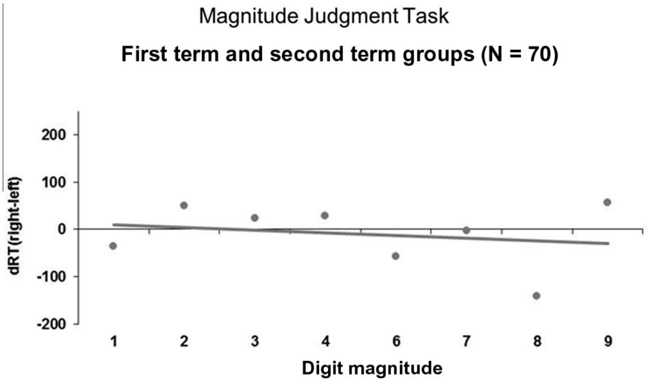

Fig. 2. Regression analysis of the difference between right- and left-hand responses using digit magnitude as a predictor variable. The resulting slope did not differ from zero; hence, no SNARC effect was present in the magnitude judgment task.

Most interesting for the current study, the interaction between response side and digit magnitude was highly significant, $F(1,77) = 10.90$, $p < .01$, $\eta^2 = .124$, revealing the presence of a SNARC effect. This interaction was further modulated by group and task order, as shown by the four-way interaction of response side, digit magnitude, group, and task order, $F(1,77) = 5.40$, $p < .05$, $\eta^2 = .07$.

To follow up on these findings, we computed a three-way mixed-model ANOVA with response side (left or right) and digit magnitude (small or large) as within-participants factors and task order (magnitude–color judgment or color–magnitude judgment) as a between-participants factor separately for each group.

In the first-term group, this analysis revealed a significant three-way interaction among all factors, $F(1,32) = 4.20$, $p < .05$, $\eta^2 = .12$, meaning that in this group the SNARC effect in the digit color judgment task differed depending on whether the color judgment task was administered first or second. No other effects reached significance (all *p*s > .05). A repeated-measures ANOVA with response side and digit magnitude as within-participants factors showed that in the first-term group a SNARC effect (interaction of response side and digit magnitude) in the color judgment task was present only in the children who were administered the magnitude judgment task before the color judgment task, $F(1,16) = 7.76$, $p < .05$, $\eta^2 = .33$, whereas in the children who were administered the color judgment first there was no interaction between those two factors, $F(1,16) = 0.04$, $p > .80$, $\eta^2 = .003$.

In the older second-term group, the three-way mixed-model ANOVA showed a main effect of response side, $F(1,45) = 8.66$, $p < .01$, $\eta^2 = .16$, due to the children being faster responding with their right hand than with their left hand (RTs = 1147.5 and 1104.7 ms for left- and right-hand responses, respectively). Interestingly, in this group composed of slightly older preschoolers, there was a significant interaction between response side and digit magnitude (i.e., SNARC effect), $F(1,45) = 10.09$, $p < .01$, $\eta^2 = .18$, and it was not modulated by task order, $F(1,45) = 0.70$, $p > .30$, $\eta^2 = .02$, meaning that the children in the second-term group showed a SNARC effect in the color judgment task, and this was not influenced by whether the task was preceded by an explicit magnitude judgment task or not.

Linear regression analysis

In the digit color judgment task, the linear regression analysis yielded an overall regression weight of −11.80. A one-tailed *t* test indicated that this slope differed significantly from 0, $t(80) = -3.50$, $p < .001$, thereby corroborating the above analysis by confirming the presence of a SNARC effect. A univariate ANOVA on regression weights with group and task order as between-participants variables yielded an interaction between group and task order, $F(1,77) = 5.40$, $p < .05$, $\eta^2 = .07$. Analyzing each group separately showed a tendency toward a main effect of order in the first-term group, $F(1,32) = 3.60$, $p < .07$, $\eta^2 = .10$, with the subgroup being administered the magnitude judgment task first displaying a significantly negative regression weight (−22.80, $p < .01$ compared with 0, one-tailed) (see Fig. 3). The subgroup being administered the digit color judgment task first did not display a significantly negative regression weight (–0.60, $p > .40$ compared with 0). In the second-term group composed of children 4 months older, task order did not play a role, $F(1,45) = 1.40$, $p > .20$, $\eta^2 = .03$. The entire group displayed a significant SNARC effect (–11.80, $p < .01$ compared with 0, one-tailed) independent of task order.

As for the magnitude judgment task, we again performed a linear regression analysis on RTs with digit magnitude as a predictor variable to further investigate the magnitude effect revealed by the ANOVA. In contrast to the magnitude judgment task, however, in the color judgment task this analysis did not reveal any significant effects (all $ps > .20$); hence, RTs did not vary linearly with digit magnitude.

In brief, preschool children attending the last year of kindergarten showed significant number–space interactions during the digit color task but not when judging digit magnitude. Moreover, in first-term children, the SNARC effect emerged only when they had performed the magnitude judgment before the color task. In line with the above-mentioned differences in numerical processing maturity, this was not the case in second-term children, who had significant SNARC effects in the color task independent of task order.

Number–space associations: Correlation analysis

We then conducted correlation analyses to evaluate how the strength of number–space interactions was related to factors reflecting the general level of numerical development. Spatio–numerical interactions were evaluated by the SNARC regression weights of the magnitude judgment and digit color judgment tasks. Factors reflecting the general numerical development were the regression weights of the distance effect in the magnitude judgment task as well as general counting abilities, counting to a maximum, and Arabic digit writing. To include a factor describing the quality of children's number representations, we also computed a difference score between the individual linear and logarithmic R^2 from the number line task (LinR^2–LogR^2). A positive score reflected a better linear fit than logarithmic fit of the estimation data and, hence, a more mature numerical representation. To consider the task strategies that children used to solve the magnitude judgment task, we determined the individual linear regression slopes of the RTs as a function of digit magnitude. Positive slopes suggested that children used a verbal counting strategy to achieve the exact magnitude comparison that was required in this task. Finally, we also considered the preschoolers' age (in months) to link the

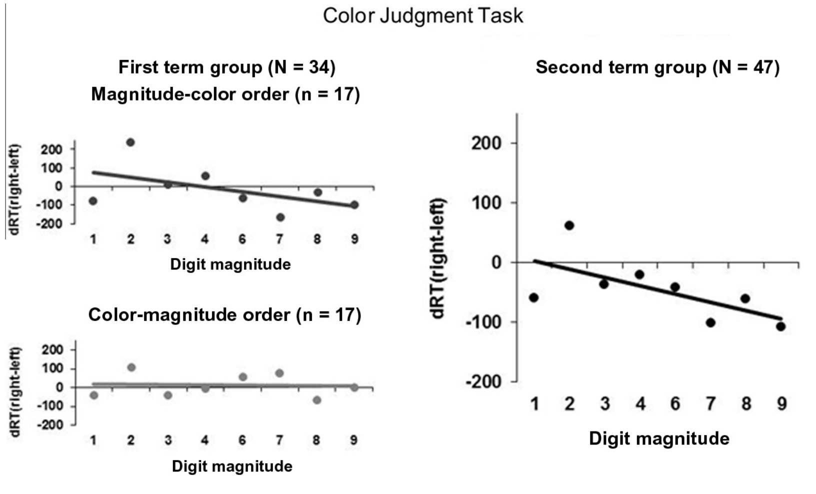

Fig. 3. Regression analysis of the two groups separately, showing a negative slope (SNARC effect) for the children in the first-term group (5.5-year-olds) who were administered the tasks in the order magnitude judgment followed by color judgment. For the second-term group (composed of 5.8-year-olds), the significantly negative slope was not modulated by order.

Table 2
Correlation table.

	1	2	3	4	5	6	7	8
1. Age								
2. SNARC Magnitude	−.03							
3. SNARC Color	.07	−.06						
4. Distance effect	−.09	−.11	−.06					
5. Counting ability	.09	−.04	.15	−.10				
6. Count to maximum	−.06	−.11	−.07	−.04	.58***			
7. Arabic Digit writing	.07	−.4**	.11	−.01	.29**	.18		
8. LinR^2–LogR^2	.17	−.12	.19	−.06	.36**	.27*	.08	
9. Counting strategy use	−.09	.30*	.07	−.20	.06	−.15	−.29*	.11

* $p < .05$.
** $p < .01$.
*** $p < .001$.

above parameters with ontogenetic development. For details of the correlation analysis, refer to Table 2. All tests are two-tailed.

The analysis revealed that the size of the SNARC effect in the magnitude judgment task was positively correlated with the recurrence to counting strategies ($r = .30$, $p < .02$). Indeed, the magnitude judgment SNARC effect tended to emerge when children no longer used counting strategies to solve simple magnitude judgments. Furthermore, the size of the SNARC effect in the magnitude judgment task was related to the Arabic writing score ($r = -.40$, $p < .01$). This result suggested that the further children knew how to write Arabic digits, the stronger their number–space associations were in a task where number magnitude was explicitly accessed. The use of counting strategies also correlated with the Arabic writing score ($r = -.30$, $p < .02$). Preschoolers who relied less on counting strategies to resolve magnitude judgments also were able to write more Arabic digits in the correct sequence. In contrast, the size of the SNARC effect obtained in the digit color judgment task did not correlate with any of the variables tested in this study. Finally, the difference score between linear and logarithmic fit values that we obtained from the number line task was positively related to counting ability and the maximum to which children could count. This means that the further children could count, the more linear their number representations were (or vice versa).

In summary, the strength of the magnitude judgment SNARC effect related positively to the children's abilities in digit writing and negatively to the use of counting strategies. In contrast, no correlations with the color judgment SNARC effect reached significance.

General discussion

The current study shows for the first time that spatio–numerical associations in the form of SNARC effects are already present when preschool children process number symbols. More specifically, children from the age of 5.5 years onward associated small Arabic digits with the left side of space and large digits with the right side of space when exact digit magnitude was irrelevant for successful task completion. Whereas the SNARC effect was significant in all children of the second-term group (composed of 5.8-year-olds), it interacted with task order in the first-term group (composed of 5.5-year-olds) such that spatio–numerical associations emerged only when the magnitude judgment task was performed before the digit color judgment task. But once activated by the explicit magnitude judgment task, the representations automatically accessed during Arabic digit processing displayed spatial properties even in first-term preschool children.

SNARC effects observed in color judgment task

During the digit color judgment task, the children did not need to access their exact numerical magnitude representation because numerical magnitude was not a task requirement. The SNARC effect that we observed during this task, however, indicates that while they were doing a simple color classification judgment on the presented digit, preschoolers concurrently activated spatially oriented numerical representations, inducing a SNARC effect. An increasing amount of evidence documents the presence of ontogenetically early tendencies to associate small/large numerical magnitude with the left/right space, respectively (for a review, see de Hevia et al., 2012). De Hevia and Spelke (2009), for instance, observed that the bisection performance of 5-year-old children was influenced by the magnitude of the nonsymbolic displays flanking the lines (see also de Hevia, 2011; Gebuis & Gevers, 2011). In addition, 4-year-old preschoolers showed spatio–numerical congruity effects when choosing the larger/smaller of two-item displays (Patro & Haman, 2012). The current study now reveals that number–space associations can even be observed when preschoolers process small Arabic digits. This observation fits with the findings that preschoolers' approximate magnitude representations reveal spatial properties (de Hevia & Spelke, 2009; Patro & Haman, 2012) and can be accessed from number symbols (Gilmore et al., 2007; Huntley-Fenner, 2001; Mejias & Schiltz, 2013). We propose that our color judgment task optimally allowed highlighting these spatial effects because the automatic activation of numerical representations was uncontaminated by intentional operations (Tzelgov & Ganor-Stern, 2005). Recent adult studies also used a similar color judgment task to successfully induce SNARC effects (Bull, Cleland, & Mitchell, 2013; Keus & Schwarz, 2005), but this methodological aspect seems critical in young children who are just acquiring formal number knowledge. To give the preschoolers sufficient time for semantic processing, digits became colored only after a small delay (i.e., 200 ms) as in the study of Mussolin and Noël (2008). We suggest that the use of this newly designed and very simple digit color judgment task also explains why we were able to observe SNARC effects much earlier than the authors of previous reports who could detect them only from 7 years onward (Berch et al., 1999; van Galen & Reitsma, 2008; White et al., 2011). Note also that these spatio–numerical effects arise considerably earlier than the automatic activations of numerical magnitude observed with numerical Stroop tasks in Western children (Girelli, Lucangeli, & Butterworth, 2000; Mussolin, 2002; Rubinsten, Henik, Berger, & Shahar-Shalev, 2002).

SNARC effects observed in magnitude judgment task

In stark contrast to the digit color judgment task, children needed to give an exact numerical answer in the magnitude comparison task and, therefore, needed to rely on some form of exact numerical representation. Whereas the SNARC effect was not significant during magnitude judgment, the significant distance effect indicates that children were solving the task similarly to what has been described in the literature (Duncan & McFarland, 1980). However, closer inspection of the graph depicting the distance effects for the first- and second-term groups shows that processing of larger numbers (i.e., 6–9) was not yet characterized by a distance effect in the younger first-term group. Moreover, correlation results indicated that especially the younger children comprising the first-term group relied on verbal counting strategies to solve the magnitude judgment instead of using more mature exact magnitude representations. In line with these observations, the first-term group did not yield significant SNARC effects during the explicit magnitude task as are typically observed in older children and adults (e.g., van Galen & Reitsma, 2008). In the older second-term group, we observed a tendency toward a SNARC effect, suggesting that more children already coactivated the spatially oriented ANS when solving the number comparison task. In line with this interpretation, the correlation data showed that the less a child relied on serial counting strategies during the magnitude judgment, the stronger his or her SNARC effect was.

The magnitude judgment SNARC effect was also strongly related to the children's ability to write Arabic digits. The further a child could write the ordinal sequence of Arabic digits without missing one number, the stronger his or her explicit SNARC effect was. In contrast, the level of verbal counting abilities did not correlate with the strength of the SNARC effect, consistent with the proposal that the list of number words (one, two, three, four, five,...) is initially learned as a meaningless ordered list (Fuson, 1988; see also Carey, 2001). The significant correlation of the magnitude judgment SNARC effect with the acquisition of number symbols (Arabic digit writing) supports the view that mastery of exact small numbers is playing a critical role in the acquisition of exact large number meanings (Carey, 2001, 2004, 2011; see also Benoit, Lehalle, Molina, Tijus, & Jouen, 2013). It has indeed been proposed that children construct a new representation of large exact numbers after they come to understand the successor function (for any known number *n* in the list, the next number is $n + 1$) and cardinal meaning (Carey, 2001, 2004; Le Corre & Carey, 2007; Noël & Rousselle, 2011; see also Carey, 2011). According to this view, children start to establish connections between the approximate ANS and the ontogenetically new exact number representation only after the construction of this new additional representation for exact large numbers is achieved (Le Corre & Carey, 2007; Noël & Rousselle, 2011). If the ability to correctly write higher digit values is a marker for more mature number representations of the associated numbers, then those children with better digit writing abilities might already have established the link between exact number symbol concepts and the underlying spatially oriented ANS. In contrast, those children with less developed digit writing abilities might still be at a previous developmental stage where this link has not yet been made. In line with this interpretation, we observed stronger magnitude judgment SNARC effects in the former children than in the latter children.

Developing numerical representations

Finally, the results from the number line task confirmed that the first-term children (5.5-year-olds) displayed a more logarithmic representation of the number range going from 0 to 20 (Siegler & Booth, 2004), indicating that they relied on the core ANS (on which magnitudes are logarithmically distributed) to solve the number line task (e.g., Dehaene, 2001; Dehaene & Cohen, 1997; Dehaene, Izard, Spelke, & Pica, 2008; Dehaene, Piazza, Pinel, & Cohen, 2003; Feigenson et al., 2004). In contrast, second-term children seemed to already possess a more mature exact numerical representation, as indicated by a better linear fit in the number line task. Because they were also less likely to use counting strategies to solve the exact magnitude judgment, we can assume that they have reached the developmental stage where they start linking the exact number representation to the spatially oriented ANS (Le Corre & Carey, 2007; Noël & Rousselle, 2011). This in turn increases the display of SNARC effects in the magnitude judgment task. This proposal is in line with the observation that the final year of preschool plays a crucial role for establishing the link between approximate magnitude and exact number symbol representations (Huntley-Fenner, 2001; Mejias & Schiltz, 2013).

To fully grasp the differential results we obtained with the color and magnitude judgment tasks, one needs to consider that they activate different types of number concepts. The color task leads to automatic activations of approximate number representations that arise as an implicit by-product of a non-numerical task. In contrast, the magnitude task relies on intentional activations of exact number concepts. From this point of view, our data suggest that in first-term children a logarithmic approximate representation of numerical magnitude was dominant. When they could rely on their ANS, which was automatically accessed from numerical symbols (i.e. Arabic digits) once this link had been activated, they showed SNARC effects (because this representation is spatially oriented and does not need to be linear to generate SNARC effects). This was the case in the digit color judgment task, which left room for an automatic activation of the ANS representation associated with number symbols (e.g., Mussolin et al., 2012; but see also Soltész et al., 2010). If, however, the first-term children were forced to produce an exact magnitude judgment, most of them relied on immature verbal strategies such as counting to solve the task and, consequently, they did not activate the ANS and did not display significant SNARC effects. With further number expertise, children then developed an exact representation of numerical magnitude that they subsequently linked to the core ANS (Le Corre & Carey, 2007; Noël & Rousselle, 2011). Through this link, exact number representations became spatially oriented, resulting in a SNARC effect also in an exact magnitude judgment task, as tended to be the case in the 5.8-year-old children of the second-term group. Our correlation data support this view because the individual SNARC effect in the magnitude judgment task, but not in the color judgment task (relying on the ontogenetically precocious ANS), was strongly related to the knowledge of Arabic digits.

Linguistic and cultural effects on developing number–space associations

The current data also contribute new elements to the theoretical discussion on the visuo–spatial versus verbal–spatial nature of the SNARC effect. Because preschool children do not yet possess a mature knowledge of the words and concepts "left" and "right" (Rigal, 1994), their SNARC effects could not be due to the fact that they associated the verbal categorical concepts small/large with left/right, respectively. Consequently, the verbal–spatial account (Gevers et al., 2010; Gevers, Verguts, Reynvoet, Caessens, & Fias, 2006; Proctor & Cho, 2006) cannot account for the spatio–numerical interactions we observed with preschool children. Because adults and already 9-year-olds show verbal coding of magnitude information (Gevers et al., 2010; Imbo, Brauwer, Fias, & Gevers, 2012), it will be interesting to investigate when the verbal coding first emerges in future studies.

Concerning the cultural origins of the left-to-right spatial component of the ANS, the results of the current study are in agreement with the finding that even before formal reading education children are influenced by a much larger general culturally determined habit of counting practice, as well as general ordering information from left to right, of which reading and writing is just one instance (Göbel, Shaki, & Fischer, 2011; Opfer & Furlong, 2011; Opfer et al., 2010; Shaki et al., 2012). By 5 years of age, most Western children indeed know to write their name (from left to right); they have seen adults or older children read from left to right, they have been shown to look at a book going from the left to the right, and they count items off going from left to right (Opfer et al., 2010).

Finally, our data also agree with the hypothesis that the determining factor in the development of number–space associations for exact numerical representations is schooling, whether it is formal or informal/parental (Gebuis, Herfs, Kenemans, de Haan, & van der Smagt, 2009; White et al., 2011). Whereas the knowledge of number symbols correlated with individual SNARC effects in magnitude judgment, age did not. Hence, the time spent in education rather than chronological age might be the determining factor (White et al., 2011).

Conclusion

The current study found a SNARC effect in preschool children as young as 5.5 years. Using an age-appropriate digit color judgment task that allowed automatic access to magnitude representation, we were able to reveal significant SNARC effects in preschoolers. Our findings indicate that from a very young age, even before entering primary school, children represent number symbols in a left-to-right fashion in Western cultures. In the magnitude judgment task, overall no significant SNARC effects were observed, but number–space associations tended to strengthen with increasing mastery of number symbols. To the best of our knowledge, the current study is the first to demonstrate the presence of left-to-right spatial associations with number symbols in preschool children.

Acknowledgments

This project was supported by a grant from the University of Luxembourg F3R-EMA-PUL-09NSP2. The first-named author is supported by an PHD-AFR grant from the Fond National de la Recherche, Luxembourg (PHD-09-160).

References

Aunola, K., Leskinen, E., Lerkkanen, M.-K., & Nurmi, J.-E. (2004). Developmental dynamics of math performance from preschool to Grade 2. *Journal of Educational Psychology, 96*, 699–713.
Barth, H. C., & Paladino, A. M. (2011). The development of numerical estimation: Evidence against a representational shift. *Developmental Science, 14*, 125–135.
Benoit, L., Lehalle, H., Molina, M., Tijus, C., & Jouen, F. (2013). Young children's mapping between arrays, number words, and digits. *Cognition, 129*, 95–101.
Berch, D. B., Foley, E. J., Hill, R. J., & Ryan, P. M. (1999). Extracting parity and magnitude from Arabic numerals: Developmental changes in number processing and mental representation. *Journal of Experimental Child Psychology, 74*, 286–308.
Booth, J. L., & Siegler, R. S. (2006). Developmental and individual differences in pure numerical estimation. *Developmental Psychology, 42*, 189–201.
Bull, R., Cleland, A. A., & Mitchell, T. (2013). Sex differences in the spatial representation of number. *Journal of Experimental Psychology: General, 142*, 181–192.
Carey, S. (2001). Cognitive foundations of arithmetic: Evolution and ontogenesis. *Mind & Language, 16*, 37–55.
Carey, S. (2004). Bootstrapping and the origin of concepts. *Daedalus, 133*, 59–68.
Carey, S. (2011). Précis of "The Origin of Concepts". *Behavioral and Brain Sciences, 34*, 113–162.
de Hevia, M. D. (2011). Sensitivity to number: Reply to Gebuis and Gevers. *Cognition, 121*, 253–255.

de Hevia, M. D., Girelli, L., & Macchi Cassia, V. (2012). Minds without language represent number through space: Origins of the mental number line. *Frontiers in Psychology, 3*, 466.
de Hevia, M. D., & Spelke, E. S. (2009). Spontaneous mapping of number and space in adults and young children. *Cognition, 110*, 198–207.
de Hevia, M. D., & Spelke, E. S. (2010). Number–space mapping in human infants. *Psychological Science, 21*, 653–660.
de Hevia, M. D., Vallar, G., & Girelli, L. (2008). Visualizing numbers in the mind's eye: The role of visuo–spatial processes in numerical abilities. *Neuroscience and Biobehavioral Reviews, 32*, 1361–1372.
Dehaene, S. (2001). Précis of "The Number Sense". *Mind & Language, 16*, 16–36.
Dehaene, S., Bossini, S., & Giraux, P. (1993). The mental representation of parity and number magnitude. *Journal of Experimental Psychology: General, 122*, 371–396.
Dehaene, S., & Cohen, L. (1997). Cerebral pathways for calculation: Double dissociation between rote verbal and quantitative knowledge of arithmetic. *Cortex, 33*, 219–250.
Dehaene, S., Dupoux, E., & Mehler, J. (1990). Is numerical comparison digital? Analogical and symbolic effects in two-digit number comparison. *Journal of Experimental Psychology: Human Perception and Performance, 16*, 626–641.
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*, 1217–1220.
Dehaene, S., Piazza, M., Pinel, P., & Cohen, L. (2003). Three parietal circuits for number processing. *Cognitive Neuropsychology, 20*, 487–506.
Duncan, E. M., & McFarland, C. E. Jr., (1980). Isolating the effects of symbolic distance and semantic congruity in comparative judgments: An additive-factors analysis. *Memory & Cognition, 8*, 612–622.
Ebersbach, M., Luwel, K., Frick, A., Onghena, P., & Verschaffel, L. (2008). The relationship between the shape of the mental number line and familiarity with numbers in 5- to 9-year old children: Evidence for a segmented linear model. *Journal of Experimental Child Psychology, 99*, 1–17.
Feigenson, L., Dehaene, S., & Spelke, E. (2004). Core systems of number. *Trends in Cognitive Sciences, 8*, 307–314.
Fias, W., Brysbaert, M., Geypens, F., & d'Ydewalle, G. (1996). The importance of magnitude information in numerical processing: Evidence from the SNARC effect. *Mathematical Cognition, 2*, 95–110.
Fias, W., & Fischer, M. H. (2005). Spatial representation of numbers. In J. I. D. Campbell (Ed.), *Handbook of mathematical cognition* (pp. 43–54). New York: Psychology Press.
Fischer, M. H., Castel, A. D., Dodd, M. D., & Pratt, J. (2003). Perceiving numbers causes spatial shifts of attention. *Nature Neuroscience, 6*, 555–556.
Fuson, K. (1988). *Children's counting and concepts of number*. New York: Springer.
Gebuis, T., & Gevers, W. (2011). Numerosities and space: Indeed a cognitive illusion! A reply to de Hevia and Spelke (2009). *Cognition, 121*, 248–255.
Gebuis, T., Herfs, I. K., Kenemans, J. L., de Haan, E. H. F., & van der Smagt, M. J. (2009). The development of automated access to symbolic and non-symbolic number knowledge in children: An ERP study. *European Journal of Neuroscience, 30*, 1999–2008.
Gevers, W., Santens, S., Dhooge, E., Chen, Q., Van den Bossche, L., Fias, W., et al (2010). Verbal–spatial and visuospatial coding of number–space interactions. *Journal of Experimental Psychology: General, 139*, 180–190.
Gevers, W., Verguts, T., Reynvoet, B., Caessens, B., & Fias, W. (2006). Numbers and space: A computational model of the SNARC effect. *Journal of Experimental Psychology: Human Perception and Performance, 32*, 32–44.
Gilmore, C. K., McCarthy, S. E., & Spelke, E. S. (2007). Symbolic arithmetic knowledge without instruction. *Nature, 447*, 589–591.
Gilmore, C. K., McCarthy, S. E., & Spelke, E. S. (2010). Non-symbolic arithmetic abilities and mathematics achievement in the first year of formal schooling. *Cognition, 115*, 394–406.
Girelli, L., Lucangeli, D., & Butterworth, B. (2000). The development of automaticity in accessing number magnitude. *Journal of Experimental Child Psychology, 76*, 104–122.
Göbel, S. M., Shaki, S., & Fischer, M. H. (2011). The cultural number line: A review of cultural and linguistic influences on the development of number processing. *Journal of Cross-Cultural Psychology, 42*, 543–565.
Huha, E. M., Berch, D. B., & Krikorian, R. (1995). Obligatory activation of magnitude information during non-numerical judgments of Arabic numerals. Poster presented at the annual meeting of the American Psychological Society, New York.
Huntley-Fenner, G. (2001). Children's understanding of number is similar to adults' and rats': Numerical estimation by 5–7-year-olds. *Cognition, 78*, B27–B40.
Imbo, I., Brauwer, J. D., Fias, W., & Gevers, W. (2012). The development of the SNARC effect: Evidence for early verbal coding. *Journal of Experimental Child Psychology, 111*, 671–680.
Keus, I. M., & Schwarz, W. (2005). Searching for the functional locus of the SNARC effect: Evidence for a response-related origin. *Memory & Cognition, 33*, 681–695.
Le Corre, M., & Carey, S. (2007). One, two, three, four, nothing more: An investigation of the conceptual sources of the verbal counting principles. *Cognition, 105*, 395–438.
Libertus, M. E., Feigenson, L., & Halberda, J. (2011). Preschool acuity of the approximate number system correlates with school math ability. *Developmental Science, 14*, 1292–1300.
Lipton, J. S., & Spelke, E. S. (2005). Preschool children's mapping of number words to nonsymbolic numerosities. *Child Development, 76*, 978–988.
Lorch, R. F., Jr., & Myers, J. L. (1990). Regression analyses of repeated measures data in cognitive research. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16*, 149–157.
Mejias, S., & Schiltz, C. (2013). Estimation abilities of large numerosities in Kindergartners. *Frontiers in Developmental Psychology, 4*, 518.
Moeller, K., Pixner, S., Kaufmann, L., & Nuerk, H.-C. (2009). Children's early mental number line: Logarithmic or decomposed linear? *Journal of Experimental Child Psychology, 103*, 503–515.
Moyer, R. S., & Landauer, T. K. (1967). Time required for judgements of numerical inequality. *Nature, 215*, 1519–1520.
Mussolin, C. (2002). *L'automaticité de l'accès à la représentation numérique chez l'enfant*. Belgium: Unpublished master's thesis, Université Catholique de Louvain.

Mussolin, C., & Noël, M.-P. (2008). Automaticity for numerical magnitude of two-digit Arabic numbers in children. *Acta Psychologica, 129*, 264–272.
Mussolin, C., Nys, J., Leybaert, J., & Content, A. (2012). Relationships between approximate number system acuity and early symbolic number abilities. *Trends in Neuroscience and Education, 1*, 21–31.
Noël, M.-P., & Rousselle, L. (2011). Developmental changes in the profiles of dyscalculia: An explanation based on a double exact-and-approximate number representation model. *Frontiers in Human Neuroscience, 5*, 165.
Opfer, J. E. (2003). Analyzing the number-line task: A tutorial. <http://www.psy.cmu.edu/ ~siegler/SiegOpfer03Tut.pdf>.
Opfer, J. E., & Furlong, E. E. (2011). How numbers bias preschoolers' spatial search. *Journal of Cross-Cultural Psychology, 42*, 682–695.
Opfer, J. E., Siegler, R. S., & Young, C. J. (2011). The powers of noise-fitting: Reply to Barth and Paladino. *Developmental Science, 14*, 1194–1206.
Opfer, J. E., Thompson, C. A., & Furlong, E. E. (2010). Early development of spatial–numeric associations: Evidence from spatial and quantitative performance of preschoolers. *Developmental Science, 13*, 761–771.
Patro, K., & Haman, M. (2012). The spatial–numerical congruity effect in preschoolers. *Journal of Experimental Child Psychology, 111*, 534–542.
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*, 33–41.
Previtali, P., de Hevia, M. D., & Girelli, L. (2010). Placing order in space: The SNARC effect in serial learning. *Experimental Brain Research, 201*, 599–605.
Proctor, R. W., & Cho, Y. S. (2006). Polarity correspondence: A general principle for performance of speeded binary classification tasks. *Psychological Bulletin, 132*, 416–442.
Restle, F. (1970). Speed of adding and comparing numbers. *Journal of Experimental Psychology, 83*, 274–278.
Rigal, R. (1994). Right–left orientation: Development of correct use of right and left terms. *Perceptual and Motor Skills, 79*, 1259–1278.
Rubinsten, O., Henik, A., Berger, A., & Shahar-Shalev, S. (2002). The development of internal representations of magnitude and their association with Arabic numerals. *Journal of Experimental Child Psychology, 81*, 74–92.
Salonen, P., Lepola, J., Vauras, M., Rauhanummi, T., Lehtinen, E., & Kinnunen, R. (1994). *Diagnostiset testit 3: Motivaatio, Metakognitio ja matematiikka*. Turku, Finland: University of Turku, Centre for Learning Research.
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.
Schneider, W., Eschmann, A., & Zuccolotto, A. (2002). *E-Prime user's guide*. Pittsburgh, PA: Psychology Software Tools.
Schweiter, M., Weinhold Zulauf, M., & von Aster, M. (2005). Die Entwicklung räumlicher Zahlenrepräsentationen und Rechenfertigkeiten bei Kindern. *Zeitschrift für Neuropsychologie, 16*, 105–113.
Sekuler, R., & Mierkiewicz, D. (1977). Children's judgments of numerical inequality. *Child Development, 48*, 630–633.
Shaki, S., Fischer, M. H., & Göbel, S. M. (2012). Direction counts: A comparative study of spatially directional counting biases in cultures with different reading directions. *Journal of Experimental Child Psychology, 112*, 275–281.
Siegler, R. S., & Booth, J. L. (2004). Development of numerical estimation in young children. *Child Development, 75*, 428–444.
Siegler, R. S., & Opfer, J. E. (2003). The development of numerical estimation: Evidence for multiple representations of numerical quantity. *Psychological Science, 14*, 237–243.
Siegler, R. S., & Robinson, M. (1982). The development of numerical understanding. *Advances in Child Development and Behavior, 16*, 241–312.
Slusser, E. B., Santiago, R. T., & Barth, H. C. (2013). Developmental change in numerical estimation. *Journal of Experimental Psychology: General, 142*, 193–208.
Soltész, F., Szucs, D., & Szucs, L. (2010). Relationships between magnitude representation, counting, and memory in 4- to 7-year-old children: A developmental study. *Behavioral and Brain Functions, 6*, 13.
Tzelgov, J., & Ganor-Stern, D. (2005). Automaticity in processing ordinal information. In J. I. D. Campbell (Ed.), *Handbook of mathematical cognition* (pp. 55–66). New York: Psychology Press.
van Dijck, J.-P., & Fias, W. (2011). A working memory account for spatial–numerical associations. *Cognition, 119*, 114–119.
van Galen, M. S., & Reitsma, P. (2008). Developing access to number magnitude: A study of the SNARC effect in 7- to 9-year-olds. *Journal of Experimental Child Psychology, 101*, 99–113.
White, S. L. J., Szűcs, D., & Soltész, F. (2011). Symbolic number: The integration of magnitude and spatial representations in children aged 6 to 8 years. *Frontiers in Psychology, 2*, 392.
Wood, G., Willmes, K., Nuerk, H.-C., & Fischer, M. H. (2008). On the cognitive link between space and number: A meta-analysis of the SNARC effect. *Psychology Science, 50*, 489–525.
Xu, F., & Spelke, E. S. (2000). Large number discrimination in 6-month-old infants. *Cognition, 74*, B1–B11.