

The knowledge of the preceding number reveals a mature understanding of the number sequence

Francesco Sella¹ & Daniela Lucangeli²

¹*Department of Psychology, University of Sheffield, UK*

²*Department of Developmental Psychology and Socialisation, University of Padova, Italy*

Keywords: symbol-grounding; approximate number system; cardinality principle; successor knowledge; predecessor knowledge; numerical cognition.

Highlights:

- CP-knowers master the successor (n+1) but not the predecessor (n-1) knowledge.
- Not all CP-knowers display a linear mapping between the ANS and the counting list.
- Predecessor knowledge and visual order of numbers related to number comparison skills.
- The knowledge of the number sequence relates to symbolic numerical magnitude.

Acknowledgements: The authors wish to thank the children and their parents for participating in the present study, as well as Giulia Alfiero and Sara Zennaro for their help in collecting data.

Introduction

One of the fascinating questions in the development of early numerical skills is understanding how children learn the numerical meaning of number words and Arabic digits, a process known as symbol-grounding (Harnad, 1990; Leibovich & Ansari, 2016). Children learn to enumerate objects in their environment by respecting three counting principles: each item in a set must be associated with only one number word (one-to-one correspondence), the counting list must be recited in the established order (stable-order principle), and the last recited number word denotes the numerosity of the set (cardinality principle; Gelman & Gallistel, 1978). The development of counting skills is widely assessed using the Give-a-number task (GaN task), in which children are presented with a large set of items (e.g., 15 tokens) and asked to give different numerical quantities. Children sequentially learn the meaning of number words from one to four and correctly give them when requested. The limit of four corresponds to the number of objects children can keep in visual short-term memory (Feigenson, Dehaene, & Spelke, 2004; Knops, Piazza, Sengupta, Eger, & Melcher, 2014; Piazza, 2010; Piazza, Fumarola, Chinello, & Melcher, 2011). At this stage, children are defined as subset-knowers because their cardinal knowledge is still limited to the first number words (Le Corre & Carey, 2007; Sarnecka & Carey, 2008). Then, children crucially understand that next counted object corresponds to the next number word in the counting list (i.e., successor function; Carey, 2004). Children become cardinal-principle knowers (CP-knowers), and they can correctly count both small (<4) and large numerical sets. The acquisition of the cardinality principle represents a crucial milestone in the development of early numerical skills: CP-knowers have understood how counting represents numbers (Gelman & Gallistel, 1978; Sarnecka, 2015). An early mastering of the cardinality principle enables children to understand the numerical meaning of numbers and constitutes an advantage for acquiring later mathematical skills (Chu, vanMarle, Rouder, & Geary, 2018; Geary et al., 2018).

The present study further explores the symbolic numerical knowledge in preschool children who were (mainly) classified as CP-knowers. In particular, we examined preschoolers' understanding of the exact numerical magnitude that is associated with number words and Arabic numbers and its relation with other numerical skills.

The ANS-to-word mapping and the later-greater principle

The mapping between numerical quantities and the counting list is usually assessed by asking individuals to estimate the number of items in briefly presented visual sets (Izard &

Dehaene, 2008; Reeve, Reynolds, Humberstone, & Butterworth, 2012). The quick presentation of the target numerosities prevents the use of serial counting and allows to assess a mapping that is not the byproduct of a procedure. CP-knowers can precisely estimate small numerosities (<3-4 items), an effect called subitizing, which depends on the parallel individuation of distinct items in the set via the Object Tracking System (OTS; Kaufman et al., 1949; Le Corre, 2014; Odic, Le Corre, & Halberda, 2015; Revkin, Piazza, Izard, Cohen, & Dehaene, 2008; Sella, Lucangeli, & Zorzi, 2018a; Trick & Pylyshyn, 1994). Large numerical quantities, instead, are processed via the Approximate Number System (ANS), in which each numerosity is represented as a Gaussian curve of activation whose width increases with numerical magnitude (Feigenson et al., 2004; Piazza, 2010). Only some CP-knowers, called mappers (i.e., CP-mappers), have established a linear mapping between the ANS and the counting list (i.e., ANS-to-word mapping) whereas some CP-knowers still lack such mapping (i.e., CP-non-mappers). Linearity is usually operationalised as a positive slope when regressing the estimates on target numerosities (Crollen, Castronovo, & Seron, 2011; Izard & Dehaene, 2008). The presence of a linear relation between estimates and target numerosities marks the understanding of the later-greater principle, that is, later number words in the counting list correspond to large numerical quantities (Le Corre, 2014; Le Corre & Carey, 2007). Children who master the cardinality principle should know that the late number words in the counting list represent large numerosities. However, this is not the case. Children acquire the later-greater principle after the cardinality principle (Le Corre & Carey, 2007). Accordingly, only CP-mappers provide estimates that increase with large numerical quantities whereas CP-non-mappers usually provide the same estimate (e.g., “five”) for large target numerosities (Odic et al., 2015). Crucially, CP-mappers have a better understanding of the symbolic system as indexed by their superior performance in comparing number words (Le Corre, 2014; but see Sella, Lucangeli, & Zorzi, 2018).

The directional property of the counting list and the order of the number line

Other studies have suggested that the ANS-to-word mapping (and more generally the ANS) has a marginal role in the development of symbolic numerical skills (Carey, 2001; Lyons, Ansari, & Beilock, 2012; Lyons, Price, Vaessen, Blomert, & Ansari, 2014; Lyons & Beilock, 2011). Conversely, number words and Arabic digits acquire their numerical meaning when children learn their reciprocal relation, in a symbol-to-symbol association (Reynvoet & Sasanguie, 2016). In this vein, the knowledge and the experience in accessing the numerical sequence play a crucial role in the symbol-grounding. The numerical sequence can be

represented in its verbal and visuospatial format, respectively, the counting list (i.e., “one”, “two”, “three”, “four”, ...) and the number line (i.e., 1-2-3-4-...).

Children early memorise the counting list and can recite it forward. However, this does not imply that children have grasped the directional property of the counting list, that is, adding one item to a set leads to the next number word (i.e., $n+1$) whereas removing one leads to the previous number word (i.e., $n-1$). Sarnecka and Carey (2008) developed the unit task, whereby children watched the experimenter adding one or two items ($n+1$ and $n+2$) to a box already containing four or five items. Children had to say the number of items in the box after the manipulation by choosing between the $n+1$ and $n+2$ answer. Similarly, in the direction task, children were presented with two sets both containing five objects, then one object was moved to one set to the other and children were requested to indicate which set had now four or six objects. CP-knowers performed the unit and direction task above the chance level as they grasp the directional property of the counting list whereas subset-knowers displayed poor performance. However, not all CP-knowers could adequately perform the two tasks (see also Davidson, Eng, & Barner, 2012). Surprisingly, all CP-knowers can proficiently add up items to a set when performing the GaN task, but not all of them can perform the same adding-one transformation in the unit task. This discrepancy might emerge from the fact some CP-knowers are implementing a rote behaviour in the GaN task without really grasping the meaning of the manipulation. In this light, the implementation of a task capable of assessing the knowledge of the successive and preceding number across the counting list without requiring any motor routine would be a perfect tool to measure the knowledge of the directional property of the counting list in preschool children. Sella and colleagues (Sella, Lucangeli, Cohen Kadosh, & Zorzi, *in press*) showed children an opaque box and told them that inside there were some felt strawberries. Then, a strawberry was either added or removed from the box, and children had to tell the number of strawberries in the box after the manipulation. The starting number of strawberries in the box varied from two to eight to explore the ability to perform $n+1$ and $n-1$ transformations in the entire numerical interval from one to nine. CP-knowers displayed an accurate performance when one item was added ($n+1$) to the box as an index of their mastering of the successor knowledge. Conversely, their accuracy in performing the $n-1$ transformation was high with small starting numbers and decreased with large starting numbers. CP-knowers display an immature mastering of the predecessor knowledge, which is successfully applied to the first, but not the later, number words in the counting list. The performance in direction task correlated with the exact symbolic numerical knowledge, as measured using a number words

comparison task, even when controlling for cardinality knowledge and memory capacity (Sella et al., *in press*).

The knowledge of the relation between numbers can also be assessed by measuring children's familiarity with the number line, as done in the number-to-position task (Berteletti, Lucangeli, Piazza, Dehaene, & Zorzi, 2010; Siegler & Opfer, 2003). In the task, children mark the position of target numbers onto a visual line entailing a numerical interval, usually from 1 to 10 for young children. Some CP-knowers, called spatial-mappers, display a linear ordering of target digits on the visual line whereas other CP-knowers, called non-spatial-mappers, tend to place all the target digits in the middle of the line or at the start or end points (Sella, Berteletti, Lucangeli, & Zorzi, 2017). CP-spatial-mappers can reproduce the spatial arrangement of numbers and demonstrate better knowledge of the exact symbolic magnitude, as measured by a higher accuracy in a digits comparison task compared to CP-non-spatial-mappers (Sella et al., 2017). To further investigate the spatial mapping of numbers in preschool children, Sella and colleagues have developed a computerised task to separately assess direction, ordinality and accuracy of the spatial mapping (DOS task; Sella, Lucangeli, & Zorzi, 2018b). In the DOS task, a digit (e.g., 2) is presented in the middle of a horizontal line and the child is requested to place an additional digit (e.g., 1) in two possible locations, one on the left side and one on the right side compared to the centred digit. The positioning of the target digit immediately determines the direction of the mapping, whether the child prefers to map numbers with increasing magnitude from left-to-right (e.g., 1-2) or right-to-left (e.g., 2-1). Then, another digit (e.g., 3) is presented, and the child is requested to place it. The positioning of the last digit reveals whether children respect the spatial order of numbers (e.g., 1-2-3) or not (e.g., 1-3-2) while also measuring the precision of the mapping, especially whether children respect the equidistance between digits (e.g., 1-2-3 vs 1-2---3). Preschool children mainly mapped numbers from left-to-right (according to the canonical direction in Western society), and the accuracy in ordering digits decreased with large triplets (e.g., 7-8-9) whereas the precision of mapping remained stable across triplets. Crucially, only spatial order related to the knowledge of the magnitude associated with digits, as measured in a digit comparison task (Sella et al., 2018b; Sella et al., *in press*). Children who can place digits in the correct spatial order within the number line can also indicate the exact numerical magnitude they represent.

Taken together, the results of the above-mentioned studies suggest that children's exact symbolic numerical knowledge might emerge from the conceptual understanding of the directional property of the counting list and the knowledge of the spatial ordinal relation

between numbers within the number line. The former represents the understanding that adding or removing items from a set is respectively associated with moving forward and backwards on the counting list. The latter indicates the knowledge of the position each number occupies within the number line. The counting list and the number line are both structures, whereby the ordinal position of a number can inform about its relative magnitude compared to the other numbers in the sequence.

The current study

In the present study, we explored the ANS-to-word mapping and the knowledge of number sequence and how these components relate to the understanding of the magnitude associated with number words and Arabic numbers. On the one hand, the cardinality principle and the ANS-to-word mapping represent a symbol-quantity relation, in which numerical quantities are mapped to symbolic numbers. On the other hand, the knowledge of the number sequence, as indexed by the directional property of the counting list and the spatial arrangement of numbers, reflect a symbol-symbol relation as numbers acquire a numerical meaning in relation to each other.

We selected a group of preschool children, who were more likely to be proficient counters (CP-knowers), to explore the knowledge of the symbolic system beyond the acquisition of the cardinality principle. We also ensured that all preschoolers in our sample knew the meaning of “numerically more” and could recite the counting list sufficiently high to perform the numerosity estimation task. Children performed a numerosity estimation task to assess the linearity of their ANS-to-word mapping. We used a direction task to measure the knowledge of the successive ($n+1$) and preceding number ($n-1$) in the counting list (Sella et al., *in press*). Children were presented with a box containing a starting number of balls, then one ball was added or removed, and children had to indicate the number of balls inside the box after the transformation. Crucially, this task does not require the implementation of rote behaviour, as the serial counting in the Give-a-number task. Moreover, we used both small and large starting numerosities, which could differentiate the knowledge of successive and preceding numbers within and outside the limit of the working memory capacity (i.e., subitizing limit in the OTS). When the box only contains 2 balls, children could perform the $n+1$ and $n-1$ transformation by tracking the number of items whereas, children had to rely on their knowledge of the directional property when the box contained a larger number of balls. We assessed the knowledge of the number line using a visual order task, whereby children saw three squares, two of them containing consecutive numbers: one in the central and one in

the right (e.g., [], [2], [3]) or left square (e.g., [1], [2], []). Children had to say the number that goes in the empty square. Compared to previously used tasks, the visual order task resembles the transformations of direction task by separately assessing what number comes after (+1) and before (-1) in the number line. Moreover, it does not require any fine motor skills, as in the number-to-position and DOS task, and the effect of lexical knowledge was removed by reading all the numbers to children. Children chose the larger between two number words and Arabic numbers to assess their knowledge of the numerical magnitude of numbers. The direct request to compare the numerical magnitude of symbols and the fact that the response cannot be achieved by implementing a behavioural routine, as in the GaN task, make the comparison tasks the perfect tools to evaluate the knowledge of the exact magnitude associated with numerical symbols. In the number comparison task, error rates and reaction times increase as a function of the numerical distance (i.e., distance effect) and the size of the numbers (i.e., size effect; (Moyer & Landauer, 1967). These effects relate to a general ratio effect as the representation of numbers overlaps with increasing numerical magnitude, as in the ANS, which is supposed to be the underlying mechanism of both non-symbolic and symbolic number comparison (Dehaene, 1992; Piazza, 2010), even though alternative models for number comparison performance have been proposed (Krajcsi, Lengyel, & Kojouharova, 2018; Verguts, Fias, & Stevens, 2005; Zorzi & Butterworth, 1999). Moreover, the speed in comparing Arabic digits has been repeatedly associated with arithmetic skills (see Schneider et al., 2016, for a review) and has been found reduced in children with developmental dyscalculia (Rousselle & Noël, 2007), which makes the exploration of the same ability in preschool children relevant, even though based on accuracy rather than response time.

Most of the studies have focused the investigation of early numerical skills on the range from 1 to 10 in preschool children. Nevertheless, it should be considered that preschool children, especially CP-knowers, should in principle correctly associate numerical quantities with the corresponding number words in their counting list, which extends above ten. In this light, preschool children might already have preliminary knowledge of the numerical magnitude represented by numbers above ten (e.g., Gilmore, McCarthy, & Spelke, 2007). Therefore, we chose to present numbers below and above ten to have a better description of children's numerical knowledge of the broader numerical sequence. We used the same target numerosities and numbers across tasks to ensure perfect comparability of the performances. Finally, we assessed verbal and visuospatial memory as domain-general control measures.

We expect some of the children in our sample to display a linear mapping between the ANS and the counting list (Le Corre, 2014; Le Corre & Carey, 2007; Sella et al., 2018a).

Children should display better performance in the $n+1$ trials of the direction task compared to the $n-1$ as children are more familiar with the process of adding up items as in the GaN task compared to remove items from a set (Sella et al., *in press*). We expect the knowledge of the spatial arrangement of numbers to decrease with large triplets (Sella et al., 2018b). Finally, we specifically examine the relative contribution of the ANS-to-word mapping, the knowledge of the counting list and the number line on number comparison skills. Such direct contrast should inform whether the knowledge of the magnitude associated with number words and Arabic numbers relates to children's knowledge of the numerical sequence or the ANS-to-word mapping.

Method

Participants

One hundred and four preschool children from five schools located in northeastern Italy took part in the study after parents, or legal guardians gave their informed consent. Parents or legal guardians also filled in a questionnaire regarding the family background and demographic information (e.g., nationality, parents' education). We excluded those children who: failed to correctly choose the larger set in the dots comparison task at least in ten out of twelve trials ($n=5$); failed to enumerate numbers at least up to 17 without committing mistakes ($n=10$). The final sample was composed of 89 children (51 boys; $M_{\text{age-months}}=70$, $SD=4$, range=60-76). All children were born in Italy. The highest level of education achieved by one of the parents/guardians indicated a middle socio-economic status for most of the families ("middle school"=28, "high school"=35, "degree/further education"=26). The study was approved by the Ethics Committee for Psychology Research of the University of Padova.

Tasks

A summary of the structures of the administered numerical tasks can be found in Table 1.

Simple dots comparison. Children indicated the larger between two numerical sets without the possibility to count. There were twelve comparisons (i.e., 10vs20, 9vs18, 15vs30, 8vs16, 9vs18, 15vs18, 12vs24, 12vs24, 15vs30, 11vs22, 14vs28, and 8vs16), which involved the same numerical ratio (i.e., 1:2). All numerical sets were composed of more than four items to prevent the use of subitizing. The numerical sets were generated using the free software Panamath (Halberda, Ly, Wilmer, Naiman, & Germine, 2012) and printed on landscape A4 papers. The two sets appeared in separated boxes on the left (yellow dots) and

right side (blue dots) of the sheet. In half of the trials, the cumulative surface area of the dots in a set was proportional to the number of dots whereas in the other half the cumulative surface area was negatively correlated with numerosity. The simple dots comparison task guaranteed that children involved in the study correctly understood the meaning of “more numerous”.

Forward enumeration. Children recited the numerical sequence starting from one. The experimenter interrupted the task when children reached 100 or when they could not go any further. Children could correct themselves immediately in case of a mistake. We recorded the individual highest recited number without committing mistakes.

Naming. Children read an Arabic digit presented on the cardboard (size: (w) 11.4 cm x (h) 9.5 cm). The experimenter showed twelve numbers in the following order: 9, 1, 17, 14, 7, 18, 12, 16, 2, 13, 3, and 8. For each child, we calculated the proportion of correct responses.

Give a Number task (GaN). We adapted this task from Wynn’s Give-a-Number (Wynn, 1990). The experimenter showed a small box with fifteen identical tokens to the child. The experimenter introduced the task as a role-playing game in which the experimenter played the role of a customer, and the child played the role of the grocer. The experimenter said: “*Let’s play the market game! You are a grocer, and I am a customer who wants to buy some lemons. Ok? Are you ready?*” The experimenter then said: “*Hello! May I have n lemon/s, please?*” As soon as the child gave the selected number of lemons, the experimenter said: “*Is this/Are these n lemon/s?*” The child could modify the number of lemons until she was sure about the number. The experimenter asked for 1, 2, 3, 4, 5, 8 and 10 lemons in order, without applying any titration. Each numerosity was asked only once. We assigned a point if the child brought the requested numerosity. For each child, we calculated the accuracy as the proportion of correct responses. We also assigned each child to a specific knower-level according to the highest peak of the posterior distribution provided by the Bayesian classification assuming the same prior probability for each knower-level (Negen, Sarnecka, & Lee, 2012).

Numerosity estimation. Children estimated the number of black squares presented on a sheet (size: (w) 19.5 cm x (h) 19.5 cm). The experimenter showed each set for approximately one second (Figure 1, panel a). There were eight target numerosities presented in the following order: 8, 17, 13, 2, 2, 13, 17, 8. Each target numerosity was presented twice:

in half of the trial the size of the squares diminished with increasing numerosity (i.e., equal cumulative surface area) whereas in the other half the size of the squares was constant (i.e., constant item size). We manipulated the size of squares to ensure that children based their estimates on the numerosity of the set rather than visual cues (Gebuis, Cohen Kadosh, & Gevers, 2016).

Direction task. The experimenter showed the child a box (size: (w) 20 cm x (h) 16 cm x (d) 30 cm), which contained some white table tennis balls (Figure 1, panel b). The experimenter said: “*Inside this box, there are n balls* [pointing at the number written on the box]!”. Then the experimenter showed the child the balls inside the box and said: “*You see, there are n balls. How many balls are in the box?*” The experimenter repeated the same question until the child correctly repeated the number of balls that were in the box. After that, the experimenter closed the box and said: “*Look carefully at what I am about to do!*” Then, depending on the type of trial, the experimenter added or removed one ball using a hole on the top of the box. Afterwards, the experimenter asked the child: “*How many balls are now in the box?*” Note that the child could not see inside the box after the manipulation. We showed eight boxes containing four starting numerosities (2, 8, 13, and 17) to which one ball was added ($n+1$) or removed ($n-1$). The eight trials followed this order of presentation: 13+1, 8-1, 17+1, 2-1, 8+1, 17-1, 2+1, and 13-1.

Visual order task (adapted from Purpura & Lonigan, 2013). The experimenter showed the child a sheet with three horizontally aligned squares, two of them containing a number (Figure 1, panel c). The experimenter said: “*Here there is the number c* [pointing to the centred one], *here there is the number $c\pm 1$* [pointing to the other number]. *Which number goes here $c\pm 1$* [pointing to the empty square]?” The experimenter always read the centred number first to avoid prompting children to mechanically reciting the number sequence (e.g., 12, 13, 14). There were four triplets of digits (1-2-3, 7-8-9, 12-13-14, 16-17-18), in which the preceding ($c-1$) or next ($c+1$) number compared to the central number was missing. The eight trials followed this order of presentation: [12][13][], [][8][9], [16][17][], [][2][3], [7][8][], [][17][18], [1][2][], and [][13][14].

Number words comparison. The experimenter named aloud the two numbers and the child had to decide which one was the larger (Figure 1, panel d). There were 12 comparisons presented twice in the following order: 1vs3, 9vs8, 12vs13, 3vs2, 18vs16, 12vs14, 1vs2, 9vs7, 16vs17, 14vs13, 2vs1, 13vs14, 18vs17, 14vs12, 3vs1, 7vs8, 16vs18, 13vs12, 8vs7, 17vs16, 7vs9, 17vs18, 2vs3, and 8vs9. The larger number in the comparison was pronounced

as first in half of the trials. The experimenter pronounced the number words with the same emphasis to prevent giving any hint about the correct answer. We computed the mean proportion of correct response for each participant.

Arabic numbers comparison. Children named or pointed at the larger between two visually presented numbers, which appeared respectively on the left and right side of the A4 paper (Figure 1, panel d). The comparisons were the same as those presented in number words comparison task. The larger number was in the right location for half of the trials. We computed the mean proportion of correct response for each participant.

Memory. Children completed the forward digit span and the forward Corsi block (BVN; Bisiacchi, Cendron, Gugliotta, Tressoldi, & Vio, 2005) to assess verbal and visuospatial memory, respectively. There were three trials for each span, starting from the span of two to seven for a total of 18 trials. The experimenter interrupted the task in case of three consecutive errors within the same span. For each child, we calculated the proportion of correct responses as the number of correct responses divided by the total number of possible trials (i.e., 18), separately for the digit span and Corsi block.

Task	Description	Details	Ability
Simple dots comparison	Choosing the larger between two numerical sets without the possibility to count.	12 comparisons. The numerosity of the sets ranged from 8 to 30. All numerical ratios were equal to $\frac{1}{2}$.	Understanding of numerically more; basic ANS acuity.
Forward enumeration	Counting up starting from 1.	Children were interrupted in case of error, when they could not count any further, or when they reached 100.	Reciting the counting list.
Naming	Naming presented numbers.	Children were requested to read numbers used in the visual order task and in the digit comparison tasks (e.g., 1, 2, 3, 7, 8, 9, 12, 13, 14, 16, 17, 18).	Transcoding from Arabic numbers to number words.

Give a number	Bringing the requested numerosity.	Seven requested numerosities: 1, 2, 3, 4, 5, 8, 10.	Cardinality principle.
Numerosity estimation	Estimating the numerosity of a set without counting.	Four target numerosities (2, 8, 13, 17), each presented twice.	ANS-to-word mapping (for target numerosities larger than 2); mastering of the later-greater principle.
Direction task	Indicating the number of items in an opaque box after one item was added (n+1) or removed (n-1) from the box.	One n+1 transformation and one n-1 transformation for each of the four starting numerosities inside the box (2, 8, 13, 17).	The knowledge of the successive (n+1) and preceding (n-1) number in the counting list; directional property of the counting list.
Visual order task	Naming the number occupying the preceding (c-1; [] [2] [3]) or successive (c+1; [1] [2] []) position on the number line.	One c-1 trial and one c+1 trial for each of the four centred numbers (2, 8, 13, 17).	Order of numbers; sequential arrangement of numbers on the number line.
Number words comparison	Choosing the larger between two number words	Twenty-four trials comparing numbers from 1 to 18.	Knowledge of exact symbolic numerical magnitude.
Arabic numbers comparison	Choosing the larger between two Arabic numbers	Twenty-four trials comparing numbers from 1 to 18.	Knowledge of exact symbolic numerical magnitude.

Table 1. Summary of the characteristics of the administered numerical tasks.

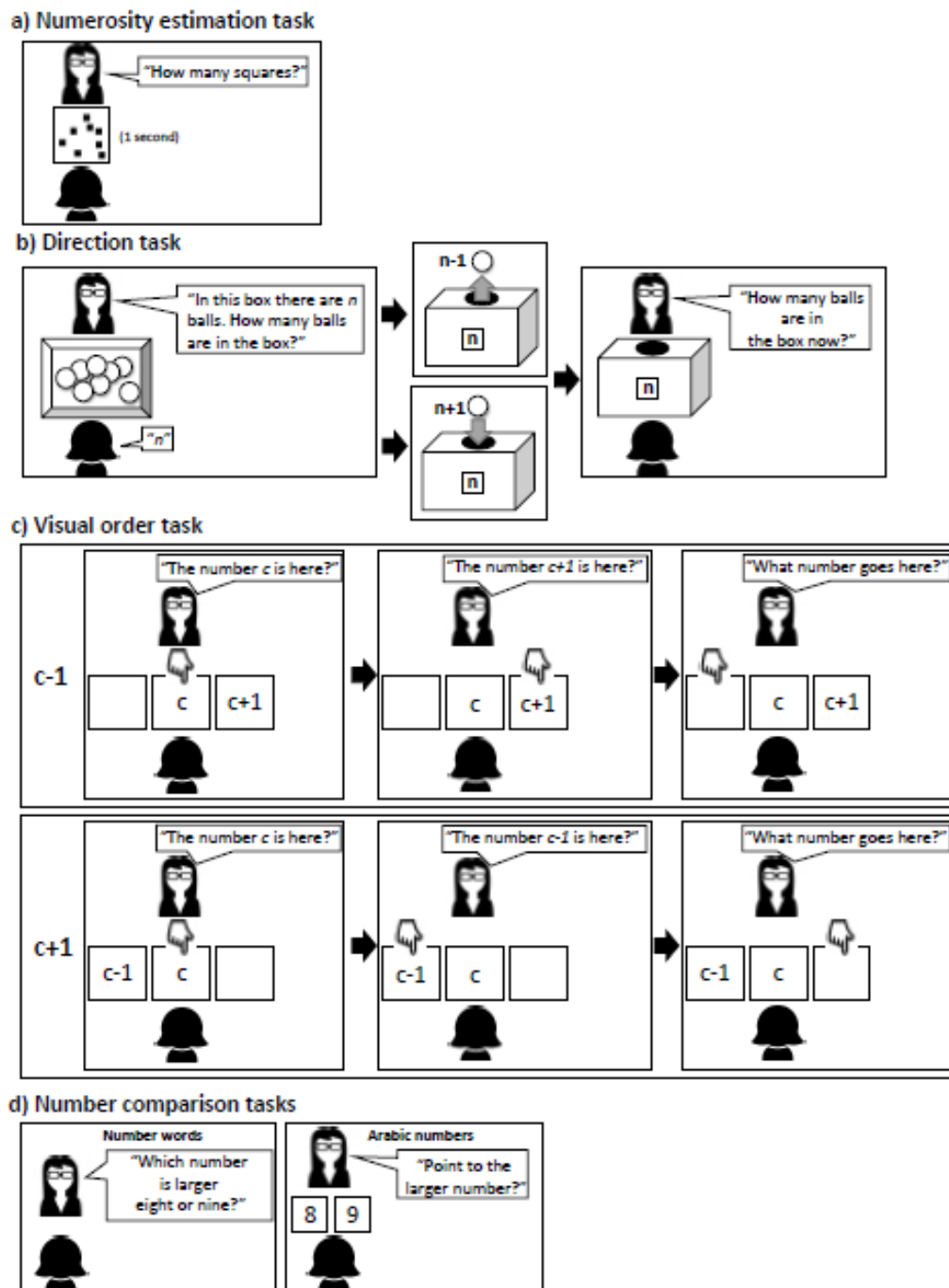


Figure 1. a) In the numerosity estimation task, the experimenter showed to the child a set composed of small squares for one second and then asked the child to estimate the number of squares. b) In the direction task, the experimenter showed the inside of a box containing some table tennis balls. Then, the experimenter closed the box and removed $(n-1)$ or added $(n+1)$ a ball into the box using the hole on the top. After the transformation, children had to say the number of table tennis balls currently inside the box. c) In the visual order task, the experimenter showed a paper sheet with three squares, two containing consecutive numbers and one empty. After pointing and naming the numbers in the squares, the experimenter asked the child to name the number which should go in the empty square. d) In the number comparison tasks, children chose the larger between two number words or two Arabic numbers. We obtained the icon of the experimenter from www.flaticon.com.

Procedure

Children completed the tasks on two sessions usually two days apart ($M=2$, $SD=1.3$, range=1-7). In the first session, children completed the following tasks in this order: GaN task, simple dots comparison, forward enumeration. The experimenter walked back in the classroom those children who did not meet the following criteria: correctly choosing the larger set in the dots comparison task at least in ten out of twelve trials and reciting forward the counting list at least up to 17 without committing mistakes. Children who met these criteria also completed the naming task, the numerosity estimation task, and the memory tasks (i.e., Corsi block and digit span) with the counterbalanced presentation of the latter. In the second session, we administered the remaining tasks (i.e., number words comparison, Arabic numbers comparison tasks, direction task and visual order task) in eight different orders. There were two different experimenters, one for first and one for the second session. The experimenter in the second session was not aware of the child's performance in the first session, thereby excluding any carryover influence across sessions.

Results

We ran the statistical analyses using the free software R (R Core Team, 2016) along with the BayesFactor package (Morey & Rouder, 2015) using default priors for Bayesian analyses. We reported Bayes factors (BF_{10}) expressing the probability of the data given H_1 relative to H_0 (i.e., values larger than 1 are in favour of H_1 whereas values smaller than 1 are in favour of H_0 ; Wagenmakers et al., 2017, 2016). When comparing regression models, we reported the Bayes factors (BF) as the ratio of BF_{s10} between compared models. If the ratio between BF_{10} of model A and BF_{10} of model B is higher larger than 1, then there is evidence for model A. Conversely, if the ratio is smaller than one there is evidence for model B. We described the evidence associated with BFs as “anecdotal” ($1/3 < BF < 3$), “moderate” ($BF < 1/3$ or $BF > 3$), “strong” ($BF < 1/10$ or $BF > 10$), “very strong” ($BF < 1/30$ or $BF > 30$), and “extreme” ($BF < 1/100$ or $BF > 100$) (Jeffreys, 1961). The comparison of BFs can quantify the evidence that data provide in favour or against a given model, or whether the results are inconclusive (Wagenmakers et al., 2018).

We calculated the reliability of the administered tasks using a bootstrap split-half approach. For each participant, we randomly selected half of the trials, and then we calculated the correlation between the two split-half performance measures 1000 times. We selected the mean of the distribution of correlation coefficients and applied the Spearman-Brown prophecy formula (i.e., $(2*r)/(1+r)$) to obtain an index of reliability. We applied this

procedure for the accuracy in the direction task ($r=0.79$), the visual order task ($r=0.77$), the naming task ($r=0.89$), the number word comparison task ($r=0.76$), the Arabic numbers comparison task ($r=0.83$), and for the absolute deviation in the numerosity estimation task ($r=0.71$). All the internal consistency values were acceptable for research purpose (Furr, 2012). We did not calculate the reliability of the Give-a-number task and the simple dots comparison task as most of the children performed these tasks at ceiling.

The raw data and the R code for the analysis can be found on the Open Science Framework (https://osf.io/xzqt4/?view_only=1424fac5dc234cff890da704ad8f8b9f).

Cardinality principle and the ANS-to-word mapping

In the GaN task, the proportion of correct responses was 0.9 ($SD=0.1$), and nine children were classified as four-knowers and eighty as CP-knowers. Most of the children mastered the cardinality principle (CP-knowers) or almost reached it (4-knowers). For the numerosity estimation task, we removed fourteen trials in which the estimates were above the 95th percentile (i.e., 70.6) of the distribution of estimates for the largest target numerosity (i.e., 17; a detailed description of the distribution of estimates can be found in the supplementary information). Then, we ran individual regression analyses with estimates as a function of target numerosities excluding the trials in which the target numerosity was two, being within the subitizing range. Then, we classified children as ANS-to-word mappers when the beta coefficient of the regression was equal to or larger than 0.3 and as non-mappers when it was lower (Le Corre & Carey, 2007). There were 40 mappers (slope: $M=1.43$, $SD=1.18$, range=0.33-5.09) and 49 non-mappers (slope: $M=0.06$, $SD=0.14$, range=-0.34-0.299). Out of 80 CP-knowers, 36 were mappers, and the remaining 44 were non-mappers.

The knowledge of the successive and preceding number

In the direction task, we examined children's ability to access a certain point in the counting list, ranging from small (e.g., 2) to large (e.g., 17) starting numbers, and perform an $n+1$ and $n-1$ transformation, which assesses the knowledge that adding an item leads to the successive number in the counting list, whereas removing one item leads to the preceding number.

We ran a Bayesian repeated measures ANOVA¹ on accuracy with starting number [2, 8, 13, 17] and transformation [n-1, n+1] as within-subjects factors (Figure 2). There was extreme evidence ($BF=5.29 \times 10^7$) in favour of the model with the interaction between starting number and transformation ($BF_{10}=3.48 \times 10^{56}$) compared to the model only including the two main effects ($BF_{10}=6.57 \times 10^{48}$). Almost all children performed the n-1 and n+1 transformation when the starting number was two. For larger starting numbers, instead, most children could perform the n+1 transformation whereas only a few successfully applied the n-1 transformation.

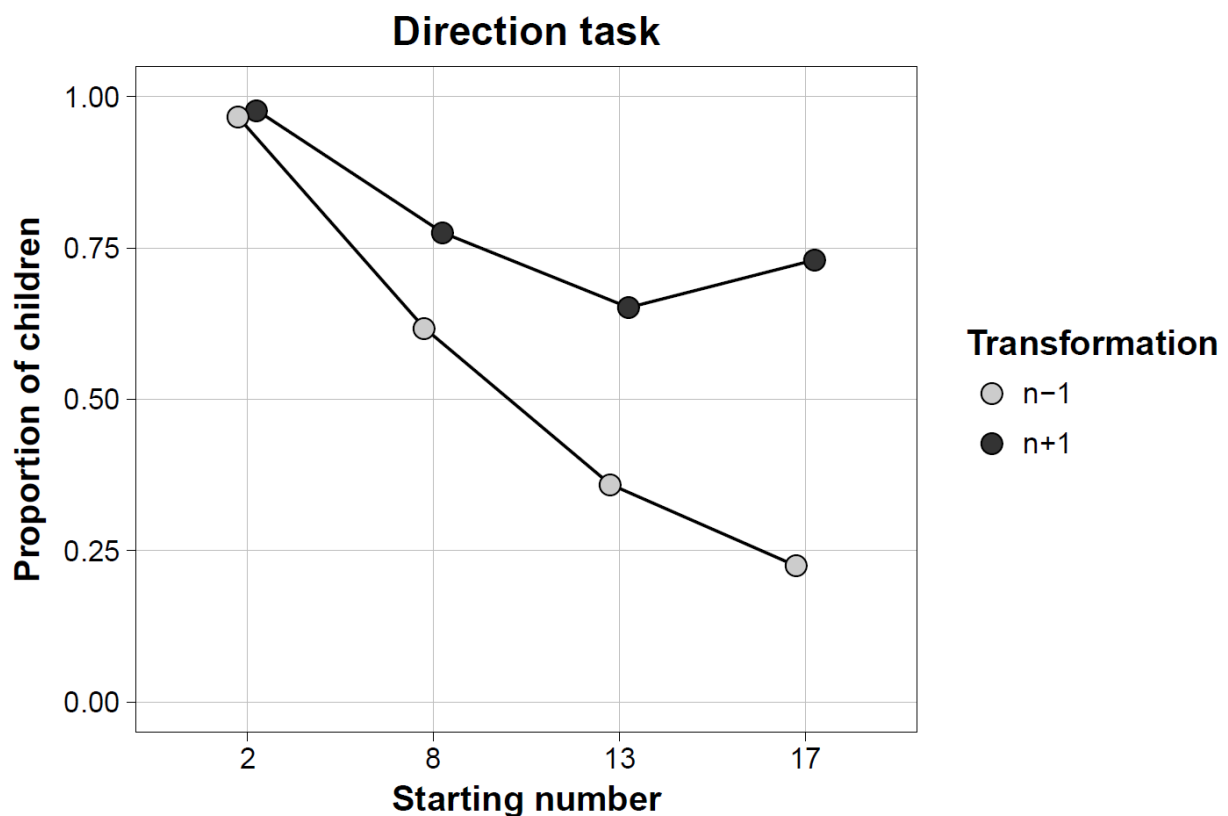


Figure 2. The proportion of children (y-axis) who gave a correct answer as a function of starting number (x-axis) separately for n-1 (light grey dots) and n+1 (dark grey dots) transformation in the direction task. The proportion of children able to perform the n+1 transformation slightly decreased when the starting number increased. The proportion of children who successfully performed the n-1 transformation dramatically decreased when the starting number increased.

We cross-tabled the proportion of correct responses of the n-1 and the n+1 trials (Table 2). Children who could perform the n-1 transformation also could perform the n+1

¹ We ran Bayesian repeated measures ANOVAs, even though the direction and the visual order task entailed binary responses (correct=1; wrong=0). The pattern of results remained the same even when we analysed the data using generalised mixed effect models (see the supplementary information).

transformation whereas the opposite was not true. Only three children were more accurate in the $n-1$ transformation than in the $n+1$ transformation.

		Accuracy of $n+1$				
		0	0.25	0.50	0.75	1
Accuracy of $n-1$	0	0	1	0	1	1
	0.25	0	8	9	3	6
	0.5	0	1	7	7	15
	0.75	0	0	1	1	11
	1	0	0	0	1	16

Table 2. Cross-table with the accuracy in the $n+1$ (columns) and $n-1$ (rows) transformations in the direction task. Values in the cells represent the number of children. Most of the children clustered on the upper right side of the matrix as indexing that children who correctly performed the $n-1$ transformations also performed $n+1$ transformations but not vice versa.

We further explored the performance in the direction task by examining the deviation between the starting number and the given response. For instance, in a $n-1$ trial with starting number eight, the child who answered “nine” gave a $n+1$ response. In Figure 3, we plotted the frequency of deviations from the starting number “ n ” separately for $n-1$ and $n+1$ trials. In the case of error when the transformation was $n-1$, children tended to respond with the $n+1$, $n+2$ or the $n-2$ number. The $n+1$ response might reflect the application of a “blind” counting forward behaviour (Barner, 2018; Carey & Barner, 2019). We speculate that the response $n+2$ might signal children’s understanding that the $n-1$ transformation was not $n+1$ and they had to say a different number, so they erroneously opted for the $n+2$ response. The $n-2$ response, instead, possibly represents a procedural error: children counted forward from one and stopped too soon at $n-2$ before getting to the right response, that is $n-1$. When the transformation was $n+1$, there were fewer errors and children tended to respond with the $n+2$, $n-1$ and $n-2$ number. The $n+2$ error might reflect the tendency to go too far in the counting list whereas the $n-2$ and $n-1$ responses might signify errors due to the alternation of $n-1$ and $n+1$ trials.

It is conceivable that the discrepancy in performance between $n+1$ and $n-1$ trials could be explained by children’s ability to inhibit the tendency to count forward in the $n-1$ trials.

Although we did not measure inhibition skills directly, age and memory skills could be considered as a proxy of executive functions and should explain the accuracy in the $n-1$ after controlling for the accuracy in the $n+1$ trials. To test this hypothesis, we ran a Bayesian regression analysis with the accuracy in the $n-1$ trials as dependent variable and the accuracy in the $n+1$ trials, age, and the performance in the digit span and Corsi block as predictors. We found strong evidence ($BF_{10}=18.73$) for the model only including the accuracy in the $n+1$ trials, with no evidence in support of the other variables. The performance in the $n-1$ trials did not relate to age and memory skills, as a proxy for executive functions, but more likely related to children's knowledge that removing one item leads to the preceding number word in the counting list.

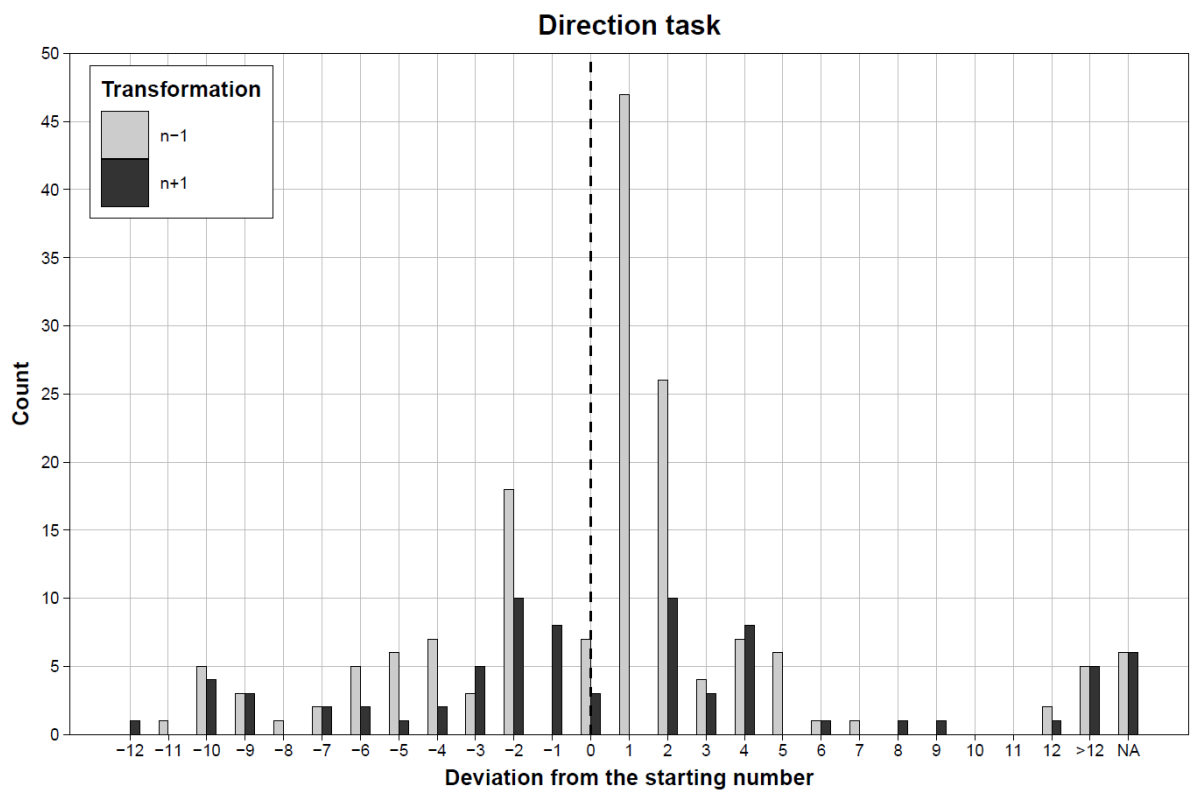


Figure 3. The frequency of deviations from the starting number (i.e., 0 on the x-axis) separately for $n-1$ (light grey bars) and $n+1$ (dark grey bars) transformations. The dashed line represents the starting number. NA (not available) denotes the number of trials in which the children responded with a non-number (e.g., “tensix”, “one and two”) or “I do not know”.

Ordinal knowledge of the number line

In the visual order task, we examined whether children knew the position occupied by a target number on the number line ranging from 2 to 17.

We ran a Bayesian repeated measures ANOVA on accuracy with centred number [2, 8, 13, 17] and target number [$c-1$, $c+1$] as within-subjects factors (Figure 4). There was

moderate evidence ($BF=7.52$) in favour of the model with the interaction between centred number and target number ($BF_{10}=4.24 \times 10^{59}$) compared to the model only including the two main effects ($BF_{10}=5.63 \times 10^{58}$).

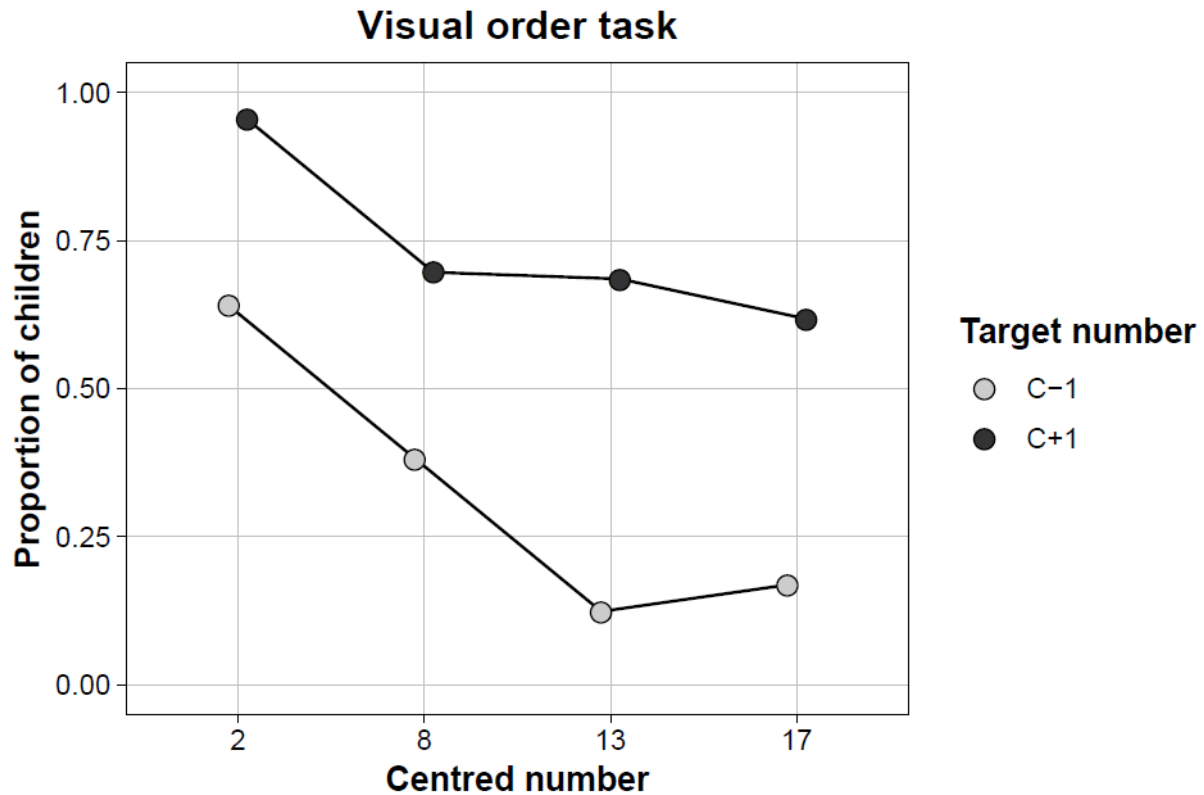


Figure 4. The proportion of children (y-axis) who gave a correct answer as a function of centred number (x-axis) separately for preceding (c-1; light grey dots) and next (c+1; dark grey dots) number in the number line.

Almost all children knew that “three” occupies the right position next to 2 whereas not all of them could tell the successive number when presented with larger numbers. Most of the children knew that “one” precedes 2, but only a few correctly named the preceding number when presented with large numbers.

We cross-tabled the proportion of correct responses of the c-1 and the c+1 trials (Table 3). Children who could perform the c-1 trials also could perform the c+1 trials whereas the opposite was not true. Only four children were more accurate in the c-1 trials than in the c+1 trials.

		Accuracy of c+1				
		0	0.25	0.50	0.75	1
Accuracy of c-1	0	3	2	7	8	10
	0.25	1	4	4	6	10
	0.5	0	4	4	3	10
	0.75	0	0	0	0	2
	1	0	0	0	0	11

Table 3. Cross-table with the accuracy in the c+1 (columns) and c-1 (rows) in the visual order task. Values in the cells represent the number of children. Most of the children clustered on the upper right side of the matrix as indexing that children who correctly performed the c-1 trials also performed n+1 trials but not vice versa.

We further explored the performance in the visual order task by examining the deviation between the centred number and the given response (Figure 5). For instance, in a n-1 trial with 8 as the centred number, a child who answered 10 gave the c+2 response. In Figure 6, we plotted the frequency of deviations from the centred number separately for preceding (c-1) and successive (c+1) target numbers. In case of error in the c+1 trials, children mainly responded with c+2 or c-2 number: the c+2 response might reflect the tendency to go too far in the number line whereas the c-2 response could represent the correct response in case it was a c-1 trial, possibly a shifting error. The strong tendency to respond with the c+2 number when the target number was the preceding one (c-1 trials) most likely reflects the application of a “blind” counting forward strategy. For instance, in case of this trial [] [13] [14], children saw and heard the numbers 13 and 14, and said 15 ignoring the instruction to name 12, that is the number that occupies the position on the left side of 13. We speculate that children lack the knowledge of the spatial arrangement of (mainly large) numbers, and then blindly did what they frequently do, that is counting forward.

It is conceivable that the discrepancy in performance between c+1 and c-1 trials could be explained by children’s ability to inhibit the tendency to count forward rather than the knowledge of the structure of the number line. We ran a Bayesian regression analysis with the accuracy in the c-1 trials as dependent variable and the accuracy in the c+1 trials, age, and the performance in the digit span and Corsi block as predictors. We found extreme evidence ($BF_{10}=5.34 \times 10^5$) for the model only including the accuracy in the c+1 trials, with no evidence in support of the other variables. The performance in the c-1 trials did not relate to age and

memory skills, as a proxy for executive functions, but more likely related to children's knowledge of the structure of the number line.

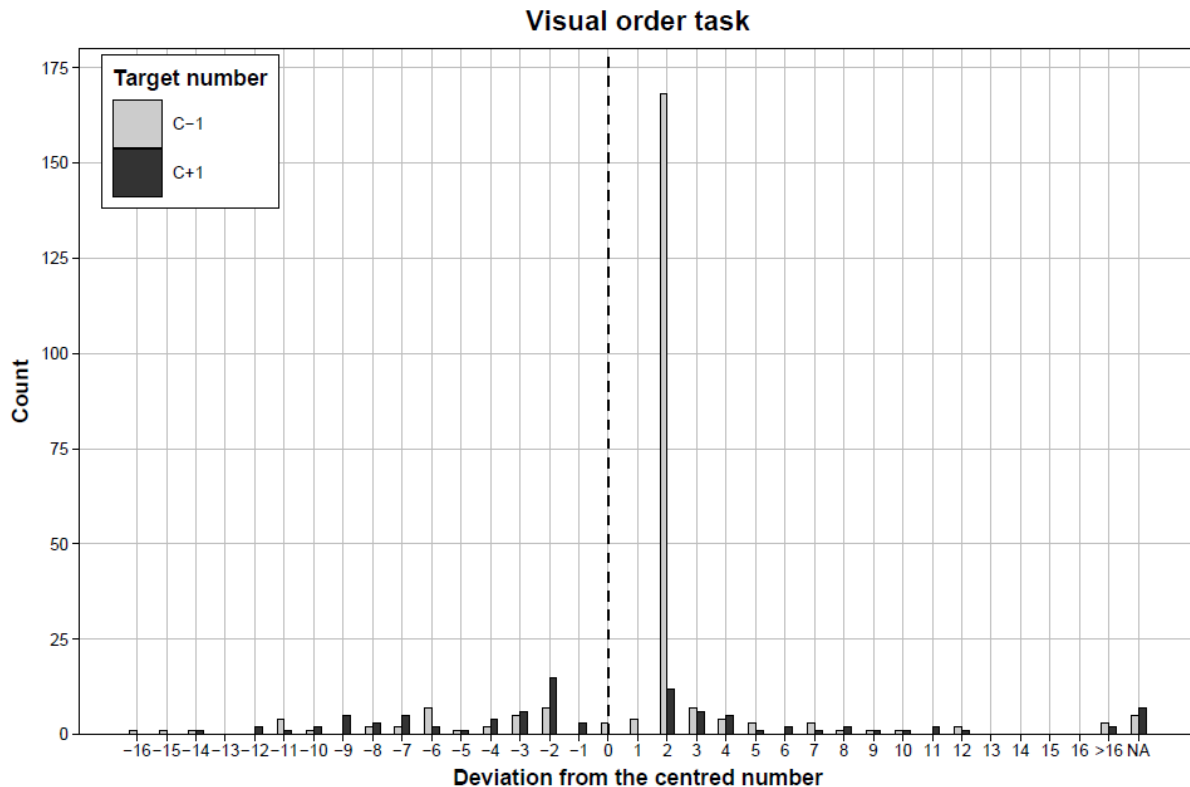


Figure 5. The frequency of deviations from the centred number (i.e., 0 on the x-axis) separately for trials in which the target number was the preceding (c-1; light grey bars) or the successive (c+1; dark grey bars). The dashed line represents the centred number. NA denotes the number of trials in which the children provide a non-number or responded: “I do not know”.

CP-non-mappers and CP-mappers

CP-non-mappers represent a remarkable stage of development to disentangle the influence of the ANS-to-word mapping on symbolic numerical skills. These children master the cardinality principle but lack a mapping between the ANS and the counting list. The lacking of the ANS-to-word mapping may prevent children from developing an understanding of the directional property of the counting list and learning the arrangement of the numbers in the number line. However, CP-non-mapper displayed a performance comparable to CP-mappers in the direction task and the visual order task (see the supplementary information for a detailed analysis). Therefore, the ANS-to-word mapping appears to have a negligible influence on children understanding of the directional property of the counting list and the knowledge of the spatial arrangement of Arabic numbers.

Correlates of exact symbolic magnitude knowledge

In the following regression analyses, we examined the relation between different numerical skills and the knowledge of the exact numerical magnitude associated with number words and Arabic numbers, as measured by number comparison tasks. In particular, we aimed to verify whether the exact symbolic magnitude knowledge is primarily related to the directional property of the counting list, the knowledge of the spatial arrangement of numbers on the number line, or the ANS-to-word mapping, in comparison with alternative models including age, generic experience with the counting list, as measured by the forward enumeration task, cardinality knowledge, domain-general factors (verbal and visuospatial memory), and also number naming in case of the Arabic numbers comparison task. We reported the descriptive statistics of the assessed variables in Table 4.

Measure	<i>M</i>	<i>SD</i>	Min	Max
Age (months)	70.04	4.11	60	76
Digit span	0.51	0.11	0.28	0.78
Corsi block	0.38	0.11	0.17	0.61
Forward enumeration (max number)	37.01	24.87	17	100
Naming	0.59	0.24	0	1
GaN	0.95	0.10	0.71	1
Direction task: n-1	0.54	0.29	0	1
Direction task: n+1	0.78	0.27	0.25	1
Visual order task: c-1	0.33	0.33	0	1
Visual order task: c+1	0.74	0.31	0	1
Number words comparison	0.73	0.16	0.46	1
Arabic number comparison	0.76	0.19	0.42	1

Table 4. Descriptive statistics. For all measures, we reported the proportion of correct responses, except for age (months) and forward enumeration (max number).

We ran a Bayesian regression analysis on accuracy in the number words comparison task (Table 5; a detailed description of the performance in the number comparison tasks can be found in the supplementary information). First, we determined the base model among the all possible models including age, digit span, Corsi block, the accuracy in the GaN task and the highest recited number in the forward enumeration task, which we divided by 100 (max score) to obtain a proportion as done for the other measures. This initial selection takes into consideration the domain-general factors, the mastering of the cardinality principle and the experience with the counting list. We applied a ranking procedure to select the domain-general model. First, we ranked all the models based on the BF_{10} , whereby larger values

entail more evidence for the model. Then, we calculated the ratio between the BF_{10} of the first model (i.e., the largest BF_{10}) and all the models with a progressive lower BF_{10} until the ratio was below 3 (i.e., anecdotal evidence). This procedure allows identifying those models with strong but similar evidence. Among these models, we selected the most parsimonious model with the largest BF_{10} and the lowest number of predictors. The selected model included the Corsi block and the highest recited number in the forward enumeration task and constituted the base model (BM=Base Model). We then compared the base model with the model also including the status of mapper² (i.e., non-mapper=0, mapper=1) in the numerosity estimation task. We found moderate evidence for the inclusion of the status of mapper (BM vs BM + Mapper; $BF=4.16$; $\Delta R^2=.05$). Then, we compared the base model with the three possible models including the accuracy of the $n-1$ and $n+1$ transformations in the direction task individually (i.e., BM + $n+1$; BM + $n-1$) and simultaneously (i.e., BM + $n+1$ + $n-1$). The inclusion of the accuracy in the $n-1$ transformation yielded the strongest evidence and was superior to the base model (BM vs BM + $n-1$; $BF=1.04 \times 10^5$, extreme evidence; $\Delta R^2=.18$). Then, we compared the base model with the three possible models including the accuracy of the $c-1$ and $c+1$ trials of the visual order task individually (i.e., BM + $c+1$; BM + $c-1$) and simultaneously (i.e., BM + $c+1$ + $c-1$). The model including the accuracy in the $c-1$ trials yielded the strongest evidence and was superior to the base model (BM vs BM + $c-1$;

² The status of mapper is a binary variable based on the individual beta coefficient of the regression analysis with estimates as a function of target numerosities in the numerosity estimation task (Mapper=beta > 0.3; Non-mapper=beta < 0.3). We also calculated an adjusted beta coefficient to have a better (continuous) measure of the ANS-to-word mapping. We adjusted the individual beta coefficient so that negative values were transformed to zero and values larger than 1 to 1. Such adjustment was necessary for two reasons. First, we did not want to penalise children with negative values as conceptually they are not worse than children with a beta of 0, they all lack the mastering of the later-greater principle (i.e., not increasing estimates associated with increasing target numerical quantities). Similarly, children with a beta larger than 1 are not conceptually better than children with a beta of 1, they all master the later-greater principle. The adjusted beta regression coefficient ranges from 0 to 1: the larger the value, the more likely the child masters the later-greater principle. When we replaced the status of mappers with the adjusted beta in the Bayesian regression analyses (Table 5 and 6), the pattern of results did not change. We found moderate evidence for a relation between the adjusted beta coefficient and number words comparison (BM vs BM + adjusted beta; $BF=3.17$) and anecdotal evidence for the absence of a relation between the adjusted beta coefficient and Arabic numbers comparison (BM vs BM + adjusted beta; $BF=0.74$).

$BF=3.53 \times 10^5$, extreme evidence; $\Delta R^2=.19$). Finally, we selected the model that provided the largest BF_{10} when all variables were taken into consideration. The selected model included the highest recited number in the forward enumeration task, the accuracy in the Corsi block, the status of mapper, the accuracy of the n-1 transformation and the c-1 trials (i.e., Best model). However, when applying the ranking procedure, the model only including the Corsi block, the accuracy in the n-1 transformation of the direction task and the c-1 trials of the visual order task emerged as the most parsimonious (i.e., Most parsimonious model). Overall, we found extreme evidence that both the mastering of the predecessor function and the ability to name the preceding number on the number line relate to the number words comparison skills. Conversely, there was moderate evidence for a relation between the ANS-to-word mapping and number words comparison skills.

	Dependent variable: Number words comparison					
	Base model (BM)	BM + Mapper	BM + Direction	BM + Visual order	Best model	Most parsimonious model
Forward enumeration	0.33 (0.22, 0.44)	0.32 (0.21, 0.42)	0.21 (0.11, 0.31)	0.16 (0.06, 0.27)	0.12 (0.02, 0.21)	
Corsi Block	0.34 (0.10, 0.59)	0.28 (0.03, 0.52)	0.23 (0.02, 0.44)	0.42 (0.21, 0.63)	0.28 (0.09, 0.47)	0.35 (0.16, 0.55)
Mapper		0.07 (0.02, 0.12)			0.04 (0.002, 0.09)	
Direction task: n-1			0.26 (0.17, 0.35)		0.18 (0.10, 0.26)	0.21 (0.12, 0.29)
Visual order task: c-1				0.24 (0.16, 0.32)	0.18 (0.09, 0.26)	0.22 (0.15, 0.29)
BF_{10}	6.29×10^6	2.62×10^7	6.53×10^{11}	2.22×10^{12}	1.81×10^{15}	6.22×10^{14}
AIC	-111.54	-116.27	-138.61	-141.29	-160	-153.57
BIC	-101.59	-103.83	-126.17	-128.85	-142.58	-141.13
R^2	0.37	0.42	0.55	0.56	0.66	0.62

Table 5. Regression models on the accuracy of the number words comparison task. Values represent regression coefficients and 95% CIs (between brackets). BF_{10} =Bayes factor against the null model; AIC= Akaike information criterion; BIC= Bayesian information criterion.

We ran a Bayesian regression analysis on accuracy in the Arabic numbers comparison task (Table 6). First, we selected the model with the strongest evidence among the all possible models including age, digit span, Corsi block, the accuracy in the GaN task, the highest recited number in the forward enumeration task (as a proportion), and the accuracy in the naming task. This initial selection took into consideration domain-general factors, the

mastering of the cardinality principle, the experience with the counting list, and the ability to read numbers as requested in the Arabic numbers comparison task. We applied the same ranking procedure and we selected the model only including the accuracy in the naming task as the base model (BM=Base Model). We then compared this base model with the model also including the status of mapper in the numerosity estimation task. We found anecdotal evidence for a relation between the status of mapper and the accuracy in the Arabic numbers comparison task (BM vs BM + Mapper; $BF=1.08$; $\Delta R^2=.03$). Then, we compared the base model with the possible three models including the accuracy of the $n-1$ and $n+1$ transformations individually and simultaneously. The model including the accuracy in the $n-1$ transformation yielded the highest evidence and was superior to the base model (BM vs BM + $n-1$; $BF=34$, very strong evidence; $\Delta R^2=.08$). Then, we compared the base model with the three possible models including the accuracy of the $c-1$ and $c+1$ trials of the visual order task individually and contemporary. The model including both the accuracy in the $c+1$ and $c-1$ trials yielded the highest evidence, but its superiority was anecdotal ($BF=1.97$) compared to a more parsimonious model only including the accuracy in the $c-1$ trials, which in turn was superior to the base model (BM vs BM + $c-1$; $BF=9.03$, moderate evidence; $\Delta R^2=.06$). Finally, we selected the model that provided the largest BF_{10} evidence when all variables were taken into consideration. The selected model included the accuracy in the naming task, in the $n-1$ transformation in the direction task and the $c-1$ trials in the visual order task (i.e., Best model). However, when applying the ranking procedure, there was anecdotal evidence for its superiority ($BF=2.6$) compared to the model only including the accuracy in the naming task and the $n-1$ transformation of the direction task (i.e., Most parsimonious model). Overall, we found respectively very strong and moderate evidence that the mastering of the predecessor knowledge and the ability to name the preceding number in the number line relate to the performance in the comparison of the Arabic numbers³. Conversely, there was

³ The direction and the visual order task each entailed only 8 trials, 4 in the $+1$ condition and 4 in the -1 condition. Therefore, the distinct contribution of the predecessor knowledge (-1 trials) compared to successor knowledge ($+1$ trials) was assessed only using a few trials. To overcome this limitation, we combined $n-1$ and $c-1$ trials into a single measure of moving backwards in the counting list and the number line and the $n+1$ and $c+1$ trials in a measure of moving forward. Then, we ran the Bayesian regression analyses reported in Table 5 and 6 replacing the mean accuracy in the $n+1$, $n-1$, $c+1$, and $c-1$ trials with the newly calculated indexes, one for moving backwards and one for moving forward. The most parsimonious model predicting the performance in the number words comparison task included the score in the Corsi block and the moving backward index while the most parsimonious model predicting the performance in the Arabic numbers comparison task included the naming accuracy and the moving backward index. These results confirm that the ability to move backwards on the number sequence relates to a mature understanding of symbolic numerical magnitude.

anecdotal evidence for a relation between the ANS-to-word mapping and Arabic numbers comparison skills.

	Dependent variable: Arabic numbers comparison				
	Base model (BM)	BM + Mapper	BM + Direction (Most parsimonious model)	BM + Visual order	Best model
Naming	0.49 (0.36, 0.62)	0.46 (0.33, 0.59)	0.30 (0.13, 0.46)	0.38 (0.24, 0.52)	0.24(0.08, 0.41)
Mapper		0.06 (0.003, 0.13)			
Direction task: n-1			0.24 (0.10, 0.38)		0.21(0.07, 0.34)
Visual order task: c-1				0.16 (0.05, 0.27)	0.13(0.02, 0.23)
BF ₁₀	1.42x10 ⁸	1.53x10 ⁸	4.86x10 ⁹	1.28x10 ⁹	1.26x10 ¹⁰
AIC	-82.75	-84.96	-92.46	-89.57	-96.41
BIC	-75.28	-75	-82.5	-79.62	-83.97
R ²	0.39	0.42	0.47	0.45	0.50

Table 6. Regression models on the accuracy of the Arabic numbers comparison task. Values represent regression coefficients and 95% CIs (between brackets). BF₁₀=Bayes factor against the null model; AIC= Akaike information criterion; BIC= Bayesian information criterion.

For all the regression models reported in Table 5 and 6: residuals were normally distributed (Shapiro tests $p > .05$); multicollinearity was absent (all variance inflation factors were lower than 2 for the models with two or more predictors); heteroscedasticity was absent (Breusch–Pagan tests $p > 0.5$, except for model “BM + Visual order” in Table 6, $p = .026$); no influential observations were found (all Cook’s distances were below or equal 1), and more generally, global assumptions were respected (Pena & Slate, 2014).

Discussion

A proficient counting represents the first step in a long process that leads children to a mature understanding of the numerical meaning of number words and Arabic numbers. In the present study, we explored how preschool children performed a variety of tasks, which specifically assessed the mastering of different numerical concepts beyond the cardinality principle. To this aim, we ensured that all children involved in the study could proficiently count, knew sufficiently well the numerical sequence and understood the meaning of numerically more.

Children displayed an almost perfect performance when estimating small numerosities (i.e., 2), whereas the variability of estimates increased with large target numerosities (Izard & Dehaene, 2008). In line with previous studies (Le Corre, 2014; Le Corre & Carey, 2007; Odic et al., 2015), only some of the CP-knowers displayed a linear mapping between the ANS and the counting list. This result supports the distinction between the acquisition of the cardinality principle and the later-greater principle (Le Corre, 2014; Le Corre & Carey, 2007). We exclude that the lack of linearity from being due to poor ANS acuity: all children reliably indicated the larger numerical set in the simple dots comparison task, whereby the ratio between numerical sets was $\frac{1}{2}$. The large numerical sets presented in the numerosity estimation task were 8, 13, and 17. Therefore, children most certainly noted that different numerical quantities were presented (i.e., 8 vs 17) and should have used different (increasing) number words when labelling target numerosities. We also exclude the possibility that children did not know the later number words in the counting list. All children could count at least up to 17 as assessed in the forward enumeration task. Two possible explanations have been proposed to explain the lack of a linear relation between estimates and target numerosities in the numerosity estimation task. On the one hand, non-mappers might find it challenging to convert the continuous ANS representation into discrete verbal labels. On the other hand, non-mappers might be able to sample among multiple verbal labels but find it difficult to select one (see Odic, Le Corre, & Halberda, 2015, for a detailed discussion).

The direction task assessed children's knowledge that adding or removing one item leads to the successive or preceding number word in the counting list. A crucial distinction should be made between those trials in which there were only two balls initially placed in the box and the remaining trials entailing larger numerosities. When the box initially contained only two balls, the number of balls varied between one and three as a result of the $n-1$ or $n+1$ transformation. Therefore, children could track each ball using the parallel individuation system and then update the number of objects kept in memory ($\bullet\bullet + \bullet = \bullet\bullet\bullet$; $\bullet\bullet - \bullet = \bullet$; Feigenson, Carey, & Hauser, 2002; Feigenson et al., 2004; Piazza, 2010). Conversely, when the box contained more than two balls, children could only rely on their knowledge of the counting list to perform the transformation. The vast majority of children could perform the $n+1$ transformation when the box initially contained two items, and only a few of them failed when the box contained more than two items. In case of wrong answers, children often went too far in the counting list responding with the $n+2$ number or responded with the $n-2$ or $n-1$ number. The $n-1$ and $n-2$ errors might be due to the switching between transformations as

children completed the task. Overall, children showed an established understanding that adding an item leads to the successive number word in the counting list (Sarnecka & Carey, 2008), at least in the proposed numerical interval (see Cheung, Rubenson, & Barner, 2017). Conversely, children in our sample displayed a limited knowledge that removing one item leads to the preceding number word (Sella et al., *in press*). They could efficiently perform the $n-1$ transformation when the box initially contained two items, but many of them failed when the box contained more items. The pattern of errors reflects a tendency to respond with the $n-2$, $n+1$, and $n+2$ number. In case of the $n-2$ errors, children may have repeated the counting list from one until the $n-1$ number, but they stopped before getting to the right response, thereby committing a procedural rather than a conceptual error. The $n+2$ errors might reflect children's intelligent guessing strategy, whereby they understand that they have to say a number different from $n+1$, which is an impossible answer given that one item was removed from the box. The $n+1$ errors, instead, may reflect children's lack of the conceptual understanding that removing an item from a set leads to the previous number word in the counting list. In this vein, the adding up might reveal children's tendency to reciting the counting list forward as a rote behaviour without fully grasping the directional structure of the counting list (Barner, 2018; Carey & Barner, 2019). We speculate that children's tendency to count forward, as they frequently did in the $n-1$ trials, possibly inflated the performance in the $n+1$ trials (Davidson et al., 2012). Some children simply named the next number in the counting list irrespective of the transformation. In this light, the accurate performance in the $n-1$ trials excludes the presence of a rote behaviour and represents a valid measure of children's understanding of the directional property of the counting list. Accordingly, children who showed high accuracy in the $n-1$ trials also showed high accuracy in the $n+1$ trials but not vice versa. Outside the limit of the parallel individuation system, most children still lack the understanding of the directional property of the counting list, despite the mastering of the cardinality principle.

In the visual order task, the vast majority of the children correctly named the number after 2 on the number line, but not all of them could name successive numbers when larger numbers were presented ($c+1$ trials). Conversely, most of the children indicated the number preceding 2, but a few of them could name larger preceding numbers ($c-1$ trials). Children had a strong tendency to respond by saying the next number in the number line as they were counting forward. We reiterate that in the $c+1$ trials the experimenter always read the centred number and then the preceding number (e.g., 13, 12) to avoid prompting children to mindlessly reciting the counting list (e.g., 12, 13). In the $c+1$ trials, counting forward leads to

the correct answer as children have to name the next number in the number line, but it leads to the wrong answer in the $c-1$ trials. The experimenter might have reinforced this tendency by first reading the centred number, and then the next number in the number line (e.g., “13 is here, 14 is here, which number goes here?”). Nevertheless, children displayed a good performance when the centred number was 2, thereby confirming their comprehension of the task instructions. Again, children’s tendency to name the next number in the counting list might have inflated the performance in the $c+1$ trials. Conversely, the $c-1$ trials required children to avoid the implementation of a rote behaviour and to know the sequential arrangement of Arabic numbers. Accordingly, children who could indicate the $c-1$ number were also able to indicate the $c+1$ number whereas the opposite was not true. The visual order task revealed that children frequently adopted a “blind” counting forward strategy, which reflects a poor knowledge of the arrangement of numbers within the number line.

Children displayed a similar level of accuracy when comparing number words and Arabic numbers with a general tendency to perform better with comparisons involving small numbers (Le Corre, 2014; Sella et al., 2018b; Sella et al., *in press*). Crucially, we found evidence that the performance in the $n-1$ trials of the direction task and the $c-1$ trials of the visual order task relate to number comparison skills compared to alternative models including cardinality knowledge, memory capacity, and also the accuracy in number naming in the case of Arabic numbers comparison (for similar results see also Sella et al., *in press*). Conversely, we found moderate/anecdotal evidence that the ANS-to-word mapping relates to the performance in the number comparison tasks. Children’s knowledge of the directional property of the counting list and the sequential arrangement of numbers on the number line strongly relates to their understanding of the magnitude associated with numerical symbols. These results have enhanced our belief that the exact symbolic magnitude knowledge may emerge from learning the numerical sequence rather than from a robust ANS-to-word mapping (Carey, Shusterman, Haward, & Distefano, 2017; Reynvoet & Sasanguie, 2016; Sella et al., 2018b). In the present study, CP-non-mappers displayed a familiarity with the counting list and the number line that was similar to CP-mappers. The lack of a linear ANS-to-word mapping did not prevent CP-non-mappers from learning the ordinal structure of the numerical sequence and improving their symbolic numerical knowledge. In this vein, number words and Arabic digits acquire numerical meaning as a consequence of their position within an ordered sequence rather than from direct association with the corresponding non-symbolic numerical quantities.

The majority of children in our sample knew that adding one item to a set leads to the next number word in the counting list. Such knowledge emerged in both the give a number task and the direction task. Similarly, most children could name the number in the right-most position in the visual order task. However, only a few children provided a correct answer when one item was removed from a set (i.e., $n-1$) in the direction task or the name of the left-most number (i.e., $c-1$) was asked in the visual order task. We speculate that the $n-1$ and $c-1$ trials were more salient in evaluating children's numerical knowledge. That is, some children have limited knowledge of the functioning of the counting list and the structure of the number line beyond small numbers. These children apply a blind counting forward strategy that leads to the correct answers in the give a number and the $n+1$ and $c+1$ trials, but reveals children's limited knowledge in the $n-1$ and $c-1$ trials. It is possible that children know that removing one item from a set leads to the preceding number word, as in the $n-1$ trials of the direction task, and can name the number occupying the left-most position, as in the $c-1$ trials visual order task, but they cannot inhibit the strong tendency to count forward. On the one hand, inhibition skills may be crucial: children know the functioning of the counting list and the structure of the number line, but struggle to accomplish the tasks because they cannot prevent the tendency of counting forward. On the other hand, numerical knowledge is more relevant: some children have limited knowledge of the functioning of the counting list and the structure of the number line. Therefore, they return to counting forward when they do not know how to accomplish the tasks. In the present study, children did not complete an inhibition task; thus, we cannot directly evaluate the influence of inhibition. However, at least three results speak in favour of the lack of knowledge rather than poor inhibition skills. First, practically all children could perform the $n-1$ transformation in the direction task when there were initially two balls in the box, thereby suggesting that children understood the instructions of the task and could inhibit the tendency to count forward. We maintain that in the case of small numbers children could perform the $n-1$ transformation by keeping the number of elements in working memory, whereas, in the case of large numbers, they had to rely on their knowledge of the directional property of the counting list. Similarly, in the case of $c-1$ trials in the visual order task, the majority of children could perform the task when the numbers were small (i.e., 1-2-3), thereby indicating that they could inhibit the tendency to answer with the consecutive number in the counting list and use the knowledge they have of the structure of the number line. One would expect the effect of inhibition to be constant across trials, irrespective of the number of items in the box or the position of the numbers on the number line. However, this is not the case. Most of the children provided the correct response in the $n-1$ and $c-1$ trials for

small numbers as they knew the answer. When children did not know the answer, they frequently return to the counting forward strategy as a consequence of a lack of knowledge rather than poor inhibition. Second, there was a strong tendency to answer with the consecutive numbers, but not always. In the direction task, in case of wrong responses, some children could inhibit the counting forward tendency and provided alternative, but still wrong, responses (e.g., $n+2$, $n-2$). In this light, the $n+2$ response might reflect an “intelligent” guessing strategy whereas the $n-2$ responses might reflect an attempt to get to the correct response by counting from one and stopping before getting to the $n-1$ (i.e., the correct answer). Admittedly, the counting forward tendency was more evident in the visual order task, whereby, in case of wrong responses, children often named the successive number (i.e., $c+2$). As mentioned before, this tendency was partially reinforced by the experimenter naming the central number and then the right-most number, thereby resembling a count forward. For example, in the case of the 7-8-9 triplet, the experimenter pointed and named the number 8 in the central position, then the number 9 in the right-most position, and finally asked the child the name of the number that goes in the left-most position. The child heard the experimenter naming aloud the numbers 8 and 9, and she completed the task by saying 10, as she was counting forward. We speculate that the counting forward strategy was stronger in the visual order task compared to the direction task because in the latter children observed a manipulation of physical items that were added or removed from a set, thereby creating a mental model of the performed transformations. Conversely, in the visual order task, there were no physical references. Children could only base their response on their knowledge of the visual order of numbers. Still, most of the children knew the sequential disposition of the numbers 1-2-3 on the number line, but lack such knowledge in the case of larger numbers. Third, inhibition skills are part of the executive functions along with working memory and increase with age (Diamond, 2014). Therefore, we decided to use age and memory skills as a proxy for inhibition. However, age and memory did not explain the performance in the $n-1$ and $c-1$ trials after controlling for the performance in the $n+1$ and $c+1$ trials, respectively. That said, we maintain that inhibition can play an important role in the acquisition of number knowledge (Bull & Scerif, 2001; Cragg & Gilmore, 2014). In particular, future studies may highlight the specific role of inhibition when different strategies are implemented in the direction and visual order task.

The correlational nature of the present study prevents to determine the directional relation between the knowledge of the number sequence, as measured by the direction and the visual order task, and number comparison skills. However, some studies have shown that

training children on spatial (i.e., number line) and the verbal order (i.e., comes after-before) improves children's early numerical skills, especially the performance in number comparison tasks (Ramani, Siegler, & Hitti, 2012; Siegler & Ramani, 2009; Xu & Lefevre, 2016; see also Merkley, Shimi, & Scerif, 2016). Future studies may disentangle the effect of the successor and predecessor knowledge on early numerical skills by training different groups of children in retrieving the successive and preceding number, respectively. One would expect the training on the predecessor knowledge to be more challenging as children are required to access the number sequence every trial, instead of simply reciting the next number as done in the training on successor knowledge. In this light, the predecessor training may be more beneficial as children will have a deep knowledge of the structure of the number sequence.

Despite the relation between the knowledge of the number sequence and number comparison skills, it is plausible that children first learn to correctly compared numbers in a trial-and-error process and, once magnitude knowledge is established, they can infer the directional property of the counting list and the relative position of numbers within the number line. This hypothesis is plausible but not parsimonious. It would imply children having extensive experience in comparing numbers, small and large ones, instead of inferring magnitude knowledge using the ordinal structure of the counting list and the number line, whereby the relative position of a number already implies its magnitude relation with other numbers within the sequence. Following this reasoning, it is plausible that number comparison, direction and order task are associated with a common latent variable, that is, the spatial representation of numbers. In this vein, the spatial disposition of numbers on the number line is a scaffolding, which entails both ordinal and magnitude information (Sella et al., 2017, 2018b; Sella et al., *in press*).

More broadly, the familiarity with the counting list and the spatial arrangement of numbers gives children a powerful frame to understand the magnitude relation between numbers. Children who have learned the numerical sequence and its directional property have achieved a mature understanding of the symbolic system. The knowledge of numerical sequence initially emerges when children compare the magnitude or have to order numbers (Sella et al., 2017, 2018a, 2018b, *in press*; Xu & Lefevre, 2016). Later in the development, the rapidity in accessing the number sequence and judging the ordinal relation between digits relates to arithmetic fluency (Lyons & Beilock, 2011; Lyons et al., 2014). The relation between ordinality and arithmetic fluency, however, is not limited to the numerical domain. The rapidity in accessing non-numerical ordered sequences, such as the alphabet, the months of the year or everyday events, also correlates with arithmetical and mathematical skills

(Helene, Sasanguie, Gevers, & Reynvoet, 2017; Morsanyi, Mahony, & McCormack, 2016; O'Connor, Morsanyi, & McCormack, 2018; Sasanguie, Lyons, De Smedt, & Reynvoet, 2017; see also Berteletti, Lucangeli, & Zorzi, 2012). Therefore, a domain-general capacity for learning and accessing ordered sequences contributes to arithmetical and mathematical achievement. Executive functions, especially the ordinal component of verbal memory, more likely support such learning (Attout & Majerus, 2014, 2017).

Conclusions

Despite the mastering of the cardinality principle, only some of the children in our sample have established a mapping between the ANS and the counting list. The direction task revealed a discrepancy between the mastering of the successor and predecessor knowledge. Children frequently said the successive number in the counting list even when one item was removed. The tendency to name the successive number might have inflated children's mastering of the successor knowledge and, more generally, the understanding of the directional property of the counting list. A strong tendency to count forward also emerged in the visual order task, especially when children were requested to name the preceding number in the number line. Such "blind" counting behaviour completely ignored the requirements of the task, thereby reflecting a poor knowledge of the sequential arrangement of large numbers. In this light, the assessment of the predecessor knowledge better revealed children's understanding of the functioning of the number sequence. Crucially, we have found strong evidence that the predecessor knowledge relates to exact symbolic magnitude knowledge, whereas we found anecdotal/moderate evidence for a relation with the ANS-to-word mapping. The results of the present study have enhanced our belief that the exact symbolic numerical magnitude emerges from a mature knowledge of the number sequence (symbol-symbol relation) rather than from a robust mapping between non-symbolic numerical quantities and the counting list (symbol-quantity relation).

References

- Attout, L., & Majerus, S. (2014). Working memory deficits in developmental dyscalculia: The importance of serial order. *Child Neuropsychology : A Journal on Normal and Abnormal Development in Childhood and Adolescence*, 21(June 2014), 1–19.
<https://doi.org/10.1080/09297049.2014.922170>
- Attout, L., & Majerus, S. (2017). Serial order working memory and numerical ordinal processing share common processes and predict arithmetic abilities. *British Journal of Developmental Psychology*, 1–14. <https://doi.org/10.1111/bjdp.12211>
- Barner, D. (2018). Numerical symbols as explanations of human perceptual experience.
<https://doi.org/https://doi.org/10.31234/osf.io/d4jca>
- Berteletti, I., Lucangeli, D., Piazza, M., Dehaene, S., & Zorzi, M. (2010). Numerical estimation in preschoolers. *Developmental Psychology*, 46(2), 545–551.
<https://doi.org/10.1037/a0017887>
- Berteletti, I., Lucangeli, D., & Zorzi, M. (2012). Representation of numerical and non-numerical order in children. *Cognition*, 124(3), 304–313.
<https://doi.org/10.1016/j.cognition.2012.05.015>
- Bisiacchi, P. S., Cendron, M., Gugliotta, M., Tressoldi, P. E., & Vio, C. (2005). *BVN 5-11: batteria di valutazione neuropsicologica per l'età evolutiva*. Trento: Centro studi Erickson.
- Carey, S. (2001). Cognitive Foundations of Arithmetic: Evolution and Ontogenesis. *Mind and Language*, 16(1), 37–55. <https://doi.org/10.1111/1468-0017.00155>
- Carey, S. (2004). Bootstrapping & the origin of concepts. *Daedalus*, 133(1), 59-68 ST-Bootstrapping & the origin of concepts. <https://doi.org/10.1162/001152604772746701>
- Carey, S., & Barner, D. (2019). Ontogenetic Origins of Human Integer Representations, 5(2), 387. https://doi.org/10.11164/jjsps.5.2_387_2
- Carey, S., Shusterman, A., Haward, P., & Distefano, R. (2017). Do analog number

- representations underlie the meanings of young children's verbal numerals? *Cognition*, 168, 243–255. <https://doi.org/https://doi.org/10.1016/j.cognition.2017.06.022>
- Cheung, P., Rubenson, M., & Barner, D. (2017). To infinity and beyond: Children generalize the successor function to all possible numbers years after learning to count. *Cognitive Psychology*, 92, 22–36. <https://doi.org/10.1016/j.cogpsych.2016.11.002>
- Chu, F. W., vanMarle, K., Rouder, J., & Geary, D. C. (2018). Children's early understanding of number predicts their later problem-solving sophistication in addition. *Journal of Experimental Child Psychology*, 169, 73–92. <https://doi.org/10.1016/j.jecp.2017.12.010>
- Crollen, V., Castronovo, J., & Seron, X. (2011). Under- and over-estimation: a bi-directional mapping process between symbolic and non-symbolic representations of number? *Experimental Psychology*, 58(1), 39–49. <https://doi.org/10.1027/1618-3169/a000064>
- Davidson, K., Eng, K., & Barner, D. (2012). Does learning to count involve a semantic induction? *Cognition*, 123(1), 162–173. <https://doi.org/10.1016/j.cognition.2011.12.013>
- Dehaene, S. (1992). Varieties Of Numerical Abilities. *Cognition*, 44, 1–42.
- Feigenson, L., Carey, S., & Hauser, M. (2002). The representations underlying infants' choice of more: object files versus analog magnitudes. *Psychological Science : A Journal of the American Psychological Society / APS*, 13(2), 150–156. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/11933999>
- Feigenson, L., Dehaene, S., & Spelke, E. (2004). Core systems of number. *Trends in Cognitive Sciences*, 8(7), 307–314. <https://doi.org/10.1016/j.tics.2004.05.002>
- Furr, M. (2012). Split-Half Reliability. In N. J. S. Book (Ed.), *Encyclopedia of Research Design* (pp. 1411–1413). Thousand Oaks {CA}: SAGE Publications, Inc.
- Geary, D. C., vanMarle, K., Chu, F. W., Rouder, J., Hoard, M. K., & Nugent, L. (2018). Early Conceptual Understanding of Cardinality Predicts Superior School-Entry Number-System Knowledge. *Psychological Science*, 29(2), 191–205. <https://doi.org/10.1177/0956797617729817>
- Gebuis, T., Cohen Kadosh, R., & Gevers, W. (2016). Sensory-integration system rather than approximate number system underlies numerosity processing : A critical review. *Acta*

- Psychologica*, 171, 1–71. <https://doi.org/10.1016/j.actpsy.2016.09.003>
- Gelman, R., & Gallistel, C. R. (1978). *The Child's Understanding of Number. Educational Researcher* (Vol. 8). Cambridge, MA: Harvard University Press.
- Gilmore, C. K., McCarthy, S. E., & Spelke, E. S. (2007). Symbolic arithmetic knowledge without instruction. *Nature*, 447(7144), 589–591. <https://doi.org/10.1038/nature05850>
- Halberda, J., Ly, R., Wilmer, J. B., Naiman, D. Q., & Germine, L. (2012). Number sense across the lifespan as revealed by a massive Internet-based sample. *Proceedings of the National Academy of Sciences of the United States of America*, 109(28), 11116–11120. <https://doi.org/10.1073/pnas.1200196109>
- Harnad, S. (1990). The symbol grounding problem. *Physica*, 42, 335–346.
- Izard, V., & Dehaene, S. (2008). Calibrating the mental number line. *Cognition*, 106(3), 1221–1247. <https://doi.org/10.1016/j.cognition.2007.06.004>
- Jeffreys, H. (1961). *Theory of probability* (3rd ed.). Oxford, UK: Oxford University Press.
- Kaufman, E. L., Lord, M. W., Reese, T. W., Volkman, J., Kaufman, E.L., Lord, M. W., ... Volkman, J. (1949). The Discrimination of Visual Number. *The American Journal of Psychology*, 62(4), 498–525.
- Knops, A., Piazza, M., Sengupta, R., Eger, E., & Melcher, D. (2014). A Shared, Flexible Neural Map Architecture Reflects Capacity Limits in Both Visual Short-Term Memory and Enumeration. *Journal of Neuroscience*, 34(30), 9857–9866. <https://doi.org/10.1523/JNEUROSCI.2758-13.2014>
- Krajcsi, A., Lengyel, G., & Kojouharova, P. (2018). Symbolic number comparison is not processed by the analog number system: Different symbolic and non-symbolic numerical distance and size effects. *Frontiers in Psychology*, 9(FEB), 1–16. <https://doi.org/10.3389/fpsyg.2018.00124>
- Le Corre, M. (2014). Children acquire the later-greater principle after the cardinal principle. *British Journal of Developmental Psychology*, 32(2), 163–177. <https://doi.org/10.1111/bjdp.12029>

- 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(2), 395–438. <https://doi.org/10.1016/j.cognition.2006.10.005>
- Leibovich, T., & Ansari, D. (2016). The symbol-grounding problem in numerical cognition: A review of theory, evidence, and outstanding questions. *Canadian Journal of Experimental Psychology*, 70(1), 12–23. <https://doi.org/10.1037/cep0000070>
- Lyons, I. M., Ansari, D., & Beilock, S. L. (2012). Symbolic estrangement: evidence against a strong association between numerical symbols and the quantities they represent. *Journal of Experimental Psychology. General*, 141(4), 635–641. <https://doi.org/10.1037/a0027248>
- Lyons, I. M., & Beilock, S. L. (2011). Numerical ordering ability mediates the relation between number-sense and arithmetic competence. *Cognition*, 121(2), 256–261. <https://doi.org/10.1016/j.cognition.2011.07.009>
- Lyons, I. M., Price, G. R., Vaessen, A., Blomert, L., & Ansari, D. (2014). Numerical predictors of arithmetic success in grades 1-6. *Developmental Science*, 5, 714–726. <https://doi.org/10.1111/desc.12152>
- Merkley, R., Shimi, A., & Scerif, G. (2016). Electrophysiological markers of newly acquired symbolic numerical representations : the role of magnitude and ordinal information. *ZDM Mathematics Education*, 48(3), 279–289. <https://doi.org/10.1007/s11858-015-0751-y>
- Morey, R. D., & Rouder, J. N. (2015). BayesFactor: Computation of Bayes Factors for Common Designs. R package version 0.9.12-2.
- Morsanyi, K., Mahony, E. O., & McCormack, T. (2016). Number comparison and number ordering as predictors of arithmetic performance in adults : Exploring the link between the two skills , and investigating the question of domain-specificity. *The Quarterly Journal of Experimental Psychology*, 0(0), 1–21. <https://doi.org/10.1080/17470218.2016.1246577>
- Moyer, R. S., & Landauer, T. K. (1967). Time required for judgements of numerical inequality. *Nature*, 215(5109), 1519–1520. <https://doi.org/10.1038/2151519a0>

- Negen, J., Sarnecka, B. W., & Lee, M. D. (2012). An Excel sheet for inferring children's number-knower levels from give-N data. *Behavior Research Methods*, 44(1), 57–66. <https://doi.org/10.3758/s13428-011-0134-4>
- O'Connor, P. A., Morsanyi, K., & McCormack, T. (2018). Young children's non-numerical ordering ability at the start of formal education longitudinally predicts their symbolic number skills and academic achievement in maths. *Developmental Science*, (October 2017), e12645. <https://doi.org/10.1111/desc.12645>
- Odic, D., Le Corre, M., & Halberda, J. (2015). Children's mappings between number words and the approximate number system. *Cognition*, 138, 102–121. <https://doi.org/10.1016/j.cognition.2015.01.008>
- Piazza, M. (2010). Neurocognitive start-up tools for symbolic number representations. *Trends in Cognitive Sciences*, 14(12), 542–551. <https://doi.org/10.1016/j.tics.2010.09.008>
- Piazza, M., Fumarola, A., Chinello, A., & Melcher, D. (2011). Subitizing reflects visuo-spatial object individuation capacity. *Cognition*, 121(1), 147–153. <https://doi.org/10.1016/j.cognition.2011.05.007>
- Purpura, D. J., & Lonigan, C. J. (2013). Informal Numeracy Skills: The Structure and Relations Among Numbering, Relations, and Arithmetic Operations in Preschool. *American Educational Research Journal*, 50(1), 178–209. <https://doi.org/10.3102/0002831212465332>
- R Core Team. (2016). R Development Core Team. *R: A Language and Environment for Statistical Computing*. Retrieved from <https://www.r-project.org/>
- 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. <https://doi.org/10.1037/a0028995>
- Reeve, R., Reynolds, F., Humberstone, J., & Butterworth, B. (2012). Stability and change in markers of core numerical competencies. *Journal of Experimental Psychology. General*, 141(4), 649–666. <https://doi.org/10.1037/a0027520>
- Revkin, S. K., Piazza, M., Izard, V., Cohen, L., & Dehaene, S. (2008). Does subitizing reflect

- numerical estimation? *Psychological Science : A Journal of the American Psychological Society / APS*, 19(6), 607–614. <https://doi.org/10.1111/j.1467-9280.2008.02130.x>
- Reynvoet, B., & Sasanguie, D. (2016). The symbol grounding problem revisited: A thorough evaluation of the ans mapping account and the proposal of an alternative account based on symbol-symbol associations. *Frontiers in Psychology*, 7(OCT), 1–11. <https://doi.org/10.3389/fpsyg.2016.01581>
- Rousselle, L., & Noël, M.-P. (2007). Basic numerical skills in children with mathematics learning disabilities: a comparison of symbolic vs non-symbolic number magnitude processing. *Cognition*, 102(3), 361–395. <https://doi.org/10.1016/j.cognition.2006.01.005>
- Sarnecka, B. W. (2015). Learning to represent exact numbers. *Synthese*. <https://doi.org/10.1007/s11229-015-0854-6>
- Sarnecka, B. W., & Carey, S. (2008). How counting represents number: what children must learn and when they learn it. *Cognition*, 108(3), 662–674. <https://doi.org/10.1016/j.cognition.2008.05.007>
- Sasanguie, D., Lyons, I. M., De Smedt, B., & Reynvoet, B. (2017). Unpacking symbolic number comparison and its relation with arithmetic in adults. *Cognition*, 165, 26–38. <https://doi.org/10.1016/j.cognition.2017.04.007>
- Schneider, M., Beeres, K., Coban, L., Merz, S., Susan Schmidt, S., Stricker, J., & De Smedt, B. (2016). Associations of non-symbolic and symbolic numerical magnitude processing with mathematical competence: A meta-analysis. *Developmental Science*, 1–16. <https://doi.org/10.1111/desc.12372>
- Sella, F., Berteletti, I., Lucangeli, D., & Zorzi, M. (2017). Preschool children use space, rather than counting, to infer the numerical magnitude of digits: evidence for a spatial mapping principle. *Cognition*, (158), 56–67.
- Sella, F., Lucangeli, D., & Zorzi, M. (2018a). Spatial and Verbal Routes to Number Comparison in Young Children. *Frontiers in Psychology*, 9(May), 1–9. <https://doi.org/10.3389/fpsyg.2018.00776>
- Sella, F., Lucangeli, D., & Zorzi, M. (2018b). Spatial order relates to the exact numerical

- magnitude of digits in young children. *Journal of Experimental Child Psychology*, 1–20.
<https://doi.org/10.1016/j.jecp.2018.09.001>
- Siegler, R. S., & Opfer, J. E. (2003). The development of numerical estimation: evidence for multiple representations of numerical quantity. *Psychological Science*, 14(3), 237–243.
- 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. <https://doi.org/10.1037/a0014239>
- Trick, L. M., & Pylyshyn, Z. W. (1994). Why Are Small and Large Numbers Enumerated Differently? A Limited-Capacity Preattentive Stage in Vision. *Psychological Review*, 101(1), 80–102.
- Verguts, T., Fias, W., & Stevens, M. (2005). A model of exact small-number representation. *Psychonomic Bulletin & Review*, 12(1), 66–80. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/15945201>
- Vos, H., Sasanguie, D., Gevers, W., & Reynvoet, B. (2017). The role of general and number-specific order processing in adults’ arithmetic performance. *Journal of Cognitive Psychology*, 5911(January). <https://doi.org/10.1080/20445911.2017.1282490>
- Wagenmakers, E.-J., Love, J., Marsman, M., Jamil, T., Ly, A., Verhagen, J., ... Morey, R. D. (2017). Bayesian inference for psychology. Part II: Example applications with JASP. *Psychonomic Bulletin and Review*, 1–19. <https://doi.org/10.3758/s13423-017-1323-7>
- Wagenmakers, E.-J., Marsman, M., Jamil, T., Ly, A., Verhagen, J., Love, J., ... Morey, R. D. (2016). Bayesian inference for psychology. Part I: Theoretical advantages and practical ramifications. *Psychonomic Bulletin & Review*, 25, 169–176.
<https://doi.org/10.3758/s13423-017-1343-3>
- Wagenmakers, E.-J., Marsman, M., Jamil, T., Ly, A., Verhagen, J., Love, J., ... Morey, R. D. (2018). Bayesian inference for psychology. Part I: Theoretical advantages and practical ramifications. *Psychonomic Bulletin and Review*, 25(1), 35–57.
<https://doi.org/10.3758/s13423-017-1343-3>
- Wynn, K. (1990). Children’s understanding of counting. *Cognition*, 36, 155–193.

- Xu, C., & Lefevre, J. (2016). Training Young Children on Sequential Relations Among Numbers and Spatial Decomposition : Differential Transfer to Number Line and Mental Transformation Tasks. *Developmental Psychology*, 52(6), 854–866.
- Zorzi, M., & Butterworth, B. (1999). A computational model of number comparison. In M. Hahn & S. Stoness (Eds.), *Twenty First Annual Meeting of the Cognitive Science Society* (p. 1). Mahwah, NJ.: Mahwah.