Making sense of number words and Arabic digits: does order count more?

Francesco Sella[1,3], Daniela Lucangeli[2], Roi Cohen Kadosh[3] & Marco Zorzi[4,5]

[1]Department of Psychology, University of Sheffield, UK
[2]Department of Developmental Psychology and Socialisation, University of Padova, Italy
[3]Department of Experimental Psychology, University of Oxford, UK
[4]Department of General Psychology and Padova Neuroscience Center, University of Padova, Italy
[5]Fondazione Ospedale and Camillo IRCCS, Venezia, Italia

This manuscript is "in press" in Child Development.

Citation: Sella, F., Lucangeli, D., Cohen Kadosh, R., & Zorzi, M. (in press). Making sense of number words and Arabic digits: does order count more? Child Development.

Acknowledgements: FS and RCK were supported by the European Research Council (Learning&Achievement 338065). MZ was supported by the CARIPARO Foundation (Excellence Projects 2017). The authors wish to thank the children and their parents for participating in the present study, as well as Chiara Monticelli for her help in collecting data.

Abstract

The ability to choose the larger between two numbers reflects a mature understanding of the magnitude associated with numerical symbols. The present study explores how the knowledge of the number sequence and memory capacity (verbal and visuospatial) relate to number comparison skills while controlling for cardinal knowledge. Preschool children's (N=140, $M_{age\text{-}in\text{-}months}$=58.9, range=41-75) mastering of the successor (n+1) and predecessor functions (n-1) as well as the spatial mapping of digits on the visual line were assessed. The ability to order digits on the visual line mediated the relation between memory capacity and number comparison skills while controlling for cardinal knowledge. Beyond cardinality, the knowledge of the (spatial) order of numbers marks the understanding of the magnitude associated with numbers.

Highlights:

- We assessed the mastering of the successor and predecessor knowledge in preschool children.
- We assessed the ability to arrange Arabic digits spatially.
- The accuracy in ordering digits relates to number comparison performance.
- The knowledge of the (spatial) order of numbers marks the understanding of the magnitude associated with numbers.

Throughout history, different cultures have developed symbolic systems to represent and manipulate numerical information (Wiese, 2003). The first stage in mastering a numerical system is learning how symbols denote numerical quantities, so that “one” refers to a set with one item, “two” refers to a set with two items and so on. Assigning an exact numerical quantity to numbers, which is a classic symbol-grounding problem (Harnad, 1990; see also Leibovich & Ansari, 2016; Reynvoet & Sasanguie, 2016), is a cornerstone in the development of numerical and mathematical skills. In Western cultures, children learn the number system by making sense of number words and Arabic digits, which are the two most-used formats to denote exact numerical information.

Numerical meaning of number words

Around the age of two, most children can recite the number words in the correct order (i.e., “one, two, three, four…”). Nevertheless, the rote declaiming of the counting list does not imply any understanding of the numerical magnitude associated with number words. At this stage, children have memorised the counting list as a string of words that stand as placeholders. Counting constitutes the first routine through which children learn to assign numerical meaning to number words. Initially, when asked to collect some objects from a large set (as in the Give-a-number task; Wynn, 1990), children grab a handful of items irrespective of the requested numerical quantity. These children are pre-number knowers because they do not know the cardinal meaning of any number words. Subsequently, children learn the numerical meaning of the number words “one”, “two”, “three” and “four” in a fixed order (Carey, 2004; Sarnecka & Carey, 2008; Wynn, 1992). These children are subset-knowers because their cardinal knowledge is limited to a portion of the counting list. The limit of four coincides with the number of objects simultaneously held in memory via the Object Tracking System (OTS; Feigenson, Dehaene, & Spelke, 2004; Piazza, 2010; Sarnecka, 2015). One proposal is that the repeated association between initial number words and small numerosities represented in the OTS permits children to achieve a conceptual leap (Carey, 2001, 2004), that is, counting one more element corresponds to the next number word in the counting list (i.e., n+1). This induction entails the extension of the cardinality principle to the entire counting list. Thus, children at this stage are defined as cardinal principle knowers (CP-knowers), even though is takes more time (approximately a couple of years after becoming CP-knowers) before they understand that the n+1 principle extends to all numbers, well beyond those in their counting list (Cheung, Rubenson, & Barner, 2017)

CP-knowers can reliably collect numerical sets corresponding to number words in their counting list, so they should know the numerical magnitude associated with number words. It follows that all CP-knowers should successfully indicate the larger between two number words (e.g., “four is more than three”, “eight is more than seven”). However, CP-knowers can successfully compare the magnitude of two number words that are smaller than 4 or when at least one number word is smaller than 4, but they fail when both number words are larger than 4 (Le Corre, 2014). This discrepancy in performance between small and large number words may emerge from the fact that children can associate small number words to the corresponding numerical quantities via OTS, thus allowing an accurate comparison. Large numerical quantities (i.e., >4), instead, are approximately represented via the Approximate Number System (ANS), in which each numerosity corresponds to a distribution of activation whose width increases with numerical magnitude (Feigenson et al., 2004; Piazza, 2010). Children who have created a mapping between the ANS and the counting list (i.e., ANS-to-word mapping) can also compare number words larger than 4 (e.g., “six” vs “ten”). These children, called CP-mappers, provide estimates that linearly increase with numerical magnitude when estimating the numerosity of large briefly presented numerical sets (Le Corre & Carey, 2007). To put it concretely, a CP-mapper may respond “six”, “seven”, and “nine” when estimating the numerosity of sets respectively containing six, eight, and ten items. Conversely, a CP-non-mapper most likely responds with the same number (e.g. “five”, “seven”) to all large target numerosities (Le Corre, 2014; Le Corre & Carey, 2007; Odic, Le Corre, & Halberda, 2015).

What does a linear ANS-to-word mapping imply for young children? A linear ANS-to-word mapping may mark the conceptual understanding that large numerical quantities correspond to number words that appear later in the counting list (i.e., later-greater principle; Le Corre, 2014). Nevertheless, some studies have shown that also subset-knowers can associate large numerical quantities to later numbers words in the counting list (Gunderson, Spaepen, & Levine, 2015; Odic, Le Corre, & Halberda, 2015; see also Barth, Starr, & Sullivan, 2009). Therefore, the later-greater principle is not an exclusive ability of CP-knowers and it does not appear to be crucial for acquiring the numerical meaning of large number words. Moreover, CP-knowers can learn to map numerical sets of ten dots to the number word “ten”, but later they cannot recognise the set with ten dots when presented with another easy to discriminate numerical set (Carey, Shusterman, Haward, & Distefano, 2017). In this vein, longitudinal studies have revealed that the classification of CP-knowers in mappers and non-mappers is unstable across time, does not correlate with ANS acuity, (Cheung, Slusser, & Shusterman, 2016; Shusterman, Slusser, Halberda, & Odic, 2016), and the children’s acquisition of cardinality contributes to symbolic numerical magnitude knowledge above and beyond any contributions of the ANS (Geary & vanMarle, 2018). These results cast some doubts on the presence of a stable mapping between the ANS and the counting list in CP-knowers. Overall, the role of the ANS-to-word mapping in the construction of an exact numerical representation of number words in young children remains unclear. Accordingly, the evidence in favour of a relation between ANS-to-word mapping and number words comparison seems to be sparse (Sella, Lucangeli, & Zorzi, 2018a, 2018b).

For instance, Sella, Lucangeli, and Zorzi (2018a) showed that mappers and non-mappers display similar accuracy when comparing small and large number words.

What does differentiate children who can compare the magnitude of number words from those who cannot? One possibility is that children acquire the exact magnitude associated with large number words when they grasp the directional property of the counting list. That is, moving forward in the counting list implies an increase in numerical magnitude whereas going backwards implies a decrease in magnitude. Sarnecka and Carey (2008; see also Dowker, 2008, for similar results) used a direction task to assess this specific understanding. The experimenter presented to children two plates, each containing five objects. The experimenter said that each plate contained five objects and then she moved one object from one plate to the other. Children indicated which plate contained six (or four) objects without having the possibility to count. Four-knowers and CP-knowers performed just above the chance level whereas other subset-knowers responded at the chance. The direction task assesses the ability to access an arbitrary point of the counting list, perform n-1 and n+1 transformations and to retrieve the next or previous number word in the counting list. Therefore, proficiency in this task marks a deep understanding of the structure of the counting list that is unlikely to reflect rote behaviour. Similarly, in the unit task (Sarnecka & Carey, 2008), children saw one or two items added to a box and then had to tell the number of objects in the box after the manipulation. Again, CP-knowers showed superior performance due to their knowledge of that adding one item leads to the next number word in the counting list. The main difference is the fact that another agent (i.e., the experimenter) adds an item to the set. It remains an open question whether children’s understanding of the directional structure of the counting list can contribute to the acquisition of number words’ meaning.

Numerical meaning of Arabic digits

While most of the research has investigated the numerical meaning of number words, only a few studies have explored how young children learn the numerical magnitude of Arabic digits. A crucial issue is whether learning the meaning of digits is a separated process or the by-product of number words knowledge and number reading skills. Children can transform a digit comparison task into a number words comparison task by reading the to-be-compared digits: in this scenario, children learn the numerical magnitude associated with number words and then extend this knowledge to digits. Conversely, learning the numerical magnitude associated with digits may follow a parallel pathway that only partially overlaps with the learning of the numerical magnitude of number words.

Some studies have explored the mappings between different numerical representations (verbal, visual, analogical) in young children. Children may create a mapping between number words and associated numerical sets, then a mapping between Arabic digits and numerical sets, and finally a mapping between number words and Arabic digits (Benoit, Lehalle, Molina, Tijus, & Jouen, 2013). Conversely, others studies have highlighted that children first map number words to numerical sets and then map number words to Arabic digits (Hurst, Anderson, & Cordes, 2016; see also, Jiménez Lira, Carver, Douglas, & LeFevre, 2017). Nevertheless, the above-mentioned studies have only focused on the mapping between different numerical representations without exploring the ability to perform number comparison, which requires ordinality or cardinality understanding. Knudsen and colleagues (Knudsen, Fischer, Henning, & Aschersleben, 2015) found that 5-years-old CP-knowers could read digits, but they were still unable to transfer their cardinality knowledge (acquired with number words) to visually presented digits. This result supports the presence of a separate (visual) route for the representation of Arabic digits, which coexists in parallel to the learning of the numerical meaning of number words. Similarly, Jiménez Lira and colleagues (2017) recently reported a detailed exploration of the mappings between number words, digits and numerical sets while assessing children's number knowledge and symbolic comparison skills. The results highlighted the specific contribution of the mapping between digits and numerical quantities to the digit comparison performance, with an indirect contribution of the word-digit and word-quantity mappings.

In a recent study, CP-knowers were classified as mappers and non-mappers based on their ability to spatially map numbers on the visual number line from 1 to 10 (i.e., number line task; Sella, Berteletti, Lucangeli, & Zorzi, 2017). CP-mappers linearly placed digits along the line whereas CP-non-mappers mainly placed digits in a non-numerical manner (e.g., all the digits in the middle of the line). Despite similar enumeration and number reading skills, CP-mappers were able to compare two visually presented digits whereas CP-non-mappers showed a poor comparison performance (for similar results see Sella, Lucangeli, & Zorzi, 2018a). A linear and accurate spatial mapping related to the exact magnitude representation of mapped numbers. Moreover, the ordinal component of the spatial mapping task correlated with the comparison of digits whereas the direction (i.e., left to right or right to left) and the precision of mapping did not (Sella et al., 2018b). A child who can spatially order a triplet of digits (e.g., 3-4-5) will most likely be able to determine the larger digit in the triplet irrespective of the direction of the spatial mapping (e.g., 3-4-5 or 5-4-3) or the accuracy in spacing digits (e.g., 3---4-5). In summary, it appears that the spatial arrangement of numbers provides a scaffold to build the magnitude representation of numbers: a digit assumes a specific numerical magnitude depending on its position on the visual line and relative to the position of other digits (spatial mapping principle; Sella et al., 2017; Sella, Lucangeli, & Zorzi, 2018a; Sella et al., 2018b). Conceivably, children have to memorise both the shape of digits and their spatial relation to achieving a correct spatial ordering. Therefore, it is plausible that the visual and spatial component of memory for object location may support the ability to map digits spatially and, in turn, the ability to compare them.

The current study

Here, we provide further evidence on the development of exact symbolic numerical representation in young children. The fact that not all CP-knowers can compare the numerical magnitude of numbers suggests that, despite the mastering of the cardinality principle in the Give-a-number task, CP-knowers still lack the knowledge of the exact numerical magnitude represented by symbolic numbers (Davidson, Eng, & Barner, 2012; Le Corre, 2014; Sella et al., 2017). In this light, the Give-a-number task is suboptimal to assess an exact symbolic numerical knowledge, which, instead, can be assessed using a number comparison task. Therefore, in this study, we use the ability to compare number words and Arabic digits as a hallmark of an exact numerical magnitude of symbolic numbers.

First, we aim to establish whether the understanding of the directional structure of the counting list can account for the magnitude representation of number words. We designed a direction task that combines the salient characteristics of the direction and unit task (Sarnecka & Carey, 2008). Children saw the experimenter adding or removing one object from an opaque box already containing some objects and had to respond by saying the number of objects in the box after the transformation. The opacity of the box prevented children from basing their responses on visual cues, such as the visible sets. Moreover, the task assessed both the knowledge that adding one item leads to the next number word in the counting list and removing one item leads to the preceding number word. We predict a strong relation between the performance in the direction task and the accuracy in the comparison of number words. In particular, the n-1 transformation may mark the understanding of the directional nature of the counting list. In this light, the n+1 transformation could be the byproduct of a rote behaviour, that is, telling the next number word in the counting list as in a mechanical forward enumeration. Conversely, the n-1 transformation is less likely to be solved as a rote behaviour. Verbal memory might support performance in the direction task, which in turn relates to the ability to compare number words. If this were the case, we would expect the performance in the direction task to mediate the relation between verbal memory and number words comparison.

Second, children may memorise the shape of Arabic digits along with their locations within the number line and use the spatial arrangement of numbers as the crucial information to derive the magnitude associated with Arabic digits (Sella et al., 2017). If this were the case, the combination of the visual and spatial components of memory for objects location may relate to learning the spatial-ordinal relation of digits, which in turn would predict Arabic digit comparison.

Third, Sella, Lucangeli, & Zorzi (2018a) found that the spatial mapping of digits was related to the comparison of Arabic digits, even when controlling for the accuracy in reading digits and the accuracy in comparing number words. Nevertheless, one cannot exclude that the observed relation between spatial mapping and symbolic magnitude could be non-spatial. Conceivably, children can use a verbal strategy (e.g., "seven should be placed here because it comes after five and six") when spatially ordering digits on the visual line. Accordingly, training the ordinality component of the counting list (e.g., "What number comes after three? What number comes before three?") improved children's performance in the number line task and a number ordering task (Xu & Lefevre, 2016). In this vein, the knowledge of the directional structure of the counting list should emerge as the crucial predictor of the comparison of both number words and Arabic digits.

Method

Participants

One hundred-seventy-one preschool children from four different schools located in northeastern of Italy took part in the study after parents, or legal guardians gave their informed consent. Parents or legal guardians also completed a questionnaire regarding the family background and demographic information (e.g., nationality, parents' education). We excluded 22 children who failed to correctly choose the larger set in the simple dots comparison task (see below) at least in ten out of twelve trials, three children who failed to enumerate numbers up to nine without committing mistakes, two children who were classified as PN-knowers, and four children who did not complete all the tasks (one child was absent on the second session; for one child the computer crashed during the number words comparison task; one child completed a different version of the digit comparison task due to the experimenter's mistake; one child did not complete the direction task). The final sample was composed of 140 children (73 boys; $M_{age\text{-}months}$=58.9, SD=9, range=41-75). Ninety-five children had both parents/guardians who were born in Italy, 15 had at least one parent who was born in Italy, and the remaining 30 had both parents who were not born in Italy. All but two children were born in Italy. The highest level of education achieved by one of the parents/guardians ("Middle school"=23, "High school"=71, "University degree"=46) indicated a middle socio-economic status for most of the families.

Tasks

Simple dots comparison. Children indicated the larger between two numerical sets without counting. There were twelve comparisons (i.e., 10vs20, 9vs18, 15vs30, 8vs16, 9vs18, 15vs18, 12vs24, 12vs24, 15vs30, 11vs22, 14vs28, and 8vs16) entailing the same numerical ratio (i.e., 1:2). All sets were larger than four to prevent the use of subitizing. Numerical sets were generated using the free software Panamath (Halberda, Ly, Wilmer, Naiman, & Germine, 2012) and presented on the computer screen. The two sets appeared in separated boxes on the left (yellow dots) and right side (blue dots) of the screen. 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 anti-correlated with numerosity. This task ensured that children understood the meaning of "more numerous", although we cannot completely rule out the possibility that children based their performance on non-numerical visual cues (Gebuis, Cohen Kadosh, & Gevers, 2016).

Forward enumeration. Children recited the numerical sequence starting from one, and the experimenter stopped them when they reached 100 or when they could not count any further. Children could correct themselves immediately if they committed a mistake. For each child, we recorded the highest recited number without committing mistakes.

Naming. Children read an Arabic digit presented on the computer screen. The experimenter showed all the digits from 1 to 9 in the following order: 3, 9, 2, 4, 7, 1, 5, 8 and 6. A child achieved one point for each correct naming, and 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 basket with fifteen identical felt strawberries 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 ga*me! You are a grocer, and I am a customer who wants to buy some delicious strawberries. Ok? Are you *ready?"* The experimenter then said: *"Hello! May* I have n strawberry/ies*, please?"* As soon as the child gave the selected number of strawberries, the experimenter said: *"Is this/Are these* n strawberry/ies*?"* The child could modify the number of strawberries until she was sure about the number. The experimenter always started asking for one strawberry and then asked for 2, 3, 4, 5, 8 and 10 strawberries in random order. The procedure was repeated twice with a brief pause between sessions. If a child failed to bring one strawberry in the first trial, then one strawberry was asked again in the second trial. The experimenter interrupted the task if a child failed both trials with one strawberry.

Memory for objects location. Children' ability to remember the visual form and the location of Arabic digits within the number line might support their ability to order numbers spatially. Therefore, we assessed children's ability to remember the form and the location of abstract stimuli in a memory for object location task. We used the Memory for Designs subtest of the NEPSY-II (Korkman, Kirk, & Kemp, 2007) in its Italian adaptation (Urgesi, Campanella, & Fabbro, 2011). Children viewed for 10 seconds a 4x4 grid with four to ten designs (abstract figures) each located in different cells. Then, the child had to select the previously seen designs from a set of cards, varying from four to 16, and place them in their original positions on the empty grid. The number of target designs and the number of cards at disposal change to progressively increase the difficulty of the task. Three- and four-year-old children completed the subtest from the first to the fourth trial whereas five- and six-years-old children started with the second trial until the fifth. The bonus score of the subtest assesses the ability to remember both the presented designs and their location within the grid. Therefore, we used the bonus score as the index of memory for object location. The maximum obtainable score varied depending on the age group according to the number of trials children completed. To overcome this issue, we divided the bonus score by the maximum score obtainable depending on the age group (i.e., 40 for 3-4 year-old and 48 for 5-6 year-old). The new rescaled bonus score almost perfectly correlated with the original raw score ($r(138)=.99$, $p<.001$), so we preferred to keep the latter as the main index of memory for object location.

Verbal Memory. We administered the Sentence Repetition subset of the NEPSY-II. The experimenter read aloud a sentence and children had to repeat it. Two points were assigned for each correctly repeated sentence, one point for repetitions with one or two errors (e.g., omission) and zero points for repetitions with more than two errors. The experimenter interrupted the subtest in case of four consecutive trials with 0 points. There were 17 sentences for a maximum of 34 points.

Direction task. In this task (Figure 1), the experimenter showed the child a box with some felt strawberries inside and said: “Inside this box, there are n strawberries! How many strawberries are in the box?” The experimenter repeated the same question until the child correctly repeated the number of strawberries that were in the box. This procedure reduced the memory load and also ensured that the children were aware of the declared number of items inside the box before the transformation. After that, the experimenter said: “Look carefully at what I am about to do!” Then, the experimenter added or removed one strawberry from the box using a hole on the top. Afterwards, the experimenter asked the child: “How many strawberries are now in the box?” After the response, the child turned her back while the experimenter pretended to add or remove items from the box. Note that the child could not see inside the box at any time, thereby the real number of objects inside the boxes did not affect the performance of the task. There were seven starting numerosities (all numbers from 2 to 8) to which the experimenter added (n+1) or removed one item (n-1). The fourteen trials followed this order of presentation: 3+1, 2-1, 4+1, 5-1, 7+1, 8-1, 6+1, 3-1, 2+1, 4-1, 5+1, 7-1, 8+1, and 6-1. The split-half reliability of accuracy was .89 (Spearman-Brown formula).

Direction, Order and Space (DOS) task (Sella et al., 2018b). In the DOS task, children arranged two sequentially presented digits spatially on a visual line (Figure 1). Children saw a black horizontal line in the middle of the screen, and a digit (hereafter centred digit) placed below the middle point of the line. Another digit (hereafter target 1) appeared above the line as soon as the child moved the mouse. According to the leftward and rightward movement of the mouse, the target 1 appeared only in two possible locations, one left and one right, equidistant (i.e., 137 pixels) from the centred digit. Children placed the target 1 by clicking one of the mouse buttons. The experimenter said: "if the number x [pointing at centred digit] is here, where should the number y [pointing at target 1] be placed?" The target 1 was always a unit more or less compared to the centred digit. Therefore, the position of target 1 determined the direction of the mapping, either left-to-right (e.g., 1-2) or right-to-left (e.g., 2-1). When the child clicked one of the mouse buttons, target 1 appeared in the selected location below the line. Then, the experimenter asked the child whether she was sure about her decision; otherwise, she could repeat the trial and place target 1 again. If the child placed target 1 exactly on the location of the centred digit, a warning message appeared informing the child that the position was already occupied and inviting the child to find another location. Once target 1 was placed, another digit (hereafter target 2) appeared above the centred digit, and children placed it on the line. The experimenter said: "If the number x [pointing at centred digit] is here and you placed the number y [pointing at target 1] here, where should the number z [pointing at target 2] be placed?" Target 2 was always a unit more or less compared to the centred digit or target 1 depending on the type of trial (see below). The placement of target 2 determined whether the child possessed a congruent spatial order of the three presented digits, regardless of directionality (e.g., 1-2-3 or 3-2-1). Moreover, the distance between the centred digit and target 1 represented a unit interval that acted as a reference to place target 2. Target 2 could be moved along the line using the mouse cursor, whose movement was restrained so that, in case of respected spatial order of the three digits, the maximum under- or over-estimation was one unit. Nonetheless, children could move the cursor of the mouse to the same extent on the opposite side of the line, even if locating target 2 in that segment would not respect ordinality. Target 2 appeared in the selected location below the line when the child clicked one of the mouse buttons. Then, the experimenter asked the child whether she was sure about her decision; otherwise, she could place target 2 again. If the child placed target 2 on the location of the centred digit or target 1, a warning message appeared inviting the child to find another location because another digit already occupied the selected one. There were three triplets 1-2-3, 4-5-6, and 7-8-9, whose digits were presented in four different orders (e.g., 2-1-3, 2-3-1, 1-2-3, and 3-2-1) twice for a total of 24 trials. Half of the trials had a two-side arrangement whereas the other half had a one-side arrangement. In the two-side trials, the target digits should be placed one on the left and one on the right side compared to the centred digit; in one-side trials, both target digits should be placed on the left or the right side of the centred digit. There was a training trial (i.e., 2-1-3) repeated twice to let the child familiarise with the task. For each trial, we recorded whether a child placed digits in the correct order. In case of a respected order, we also measured the direction of mapping (i.e., left-to-right or right-to-left) and the absolute distance (absolute error) in pixels between the estimated and the correct position of the target 2. In the case of correct ordinality, the absolute error could vary between 0 and 1 given the constraint to the movements of the mouse. The reader can find a video illustrating the DOS task at https://osf.io/gczpw/?view_only=df80d2a1e5984e008eb8a11cf8860b32. The split-half reliability was .83 for ordinality and .88 for absolute error.

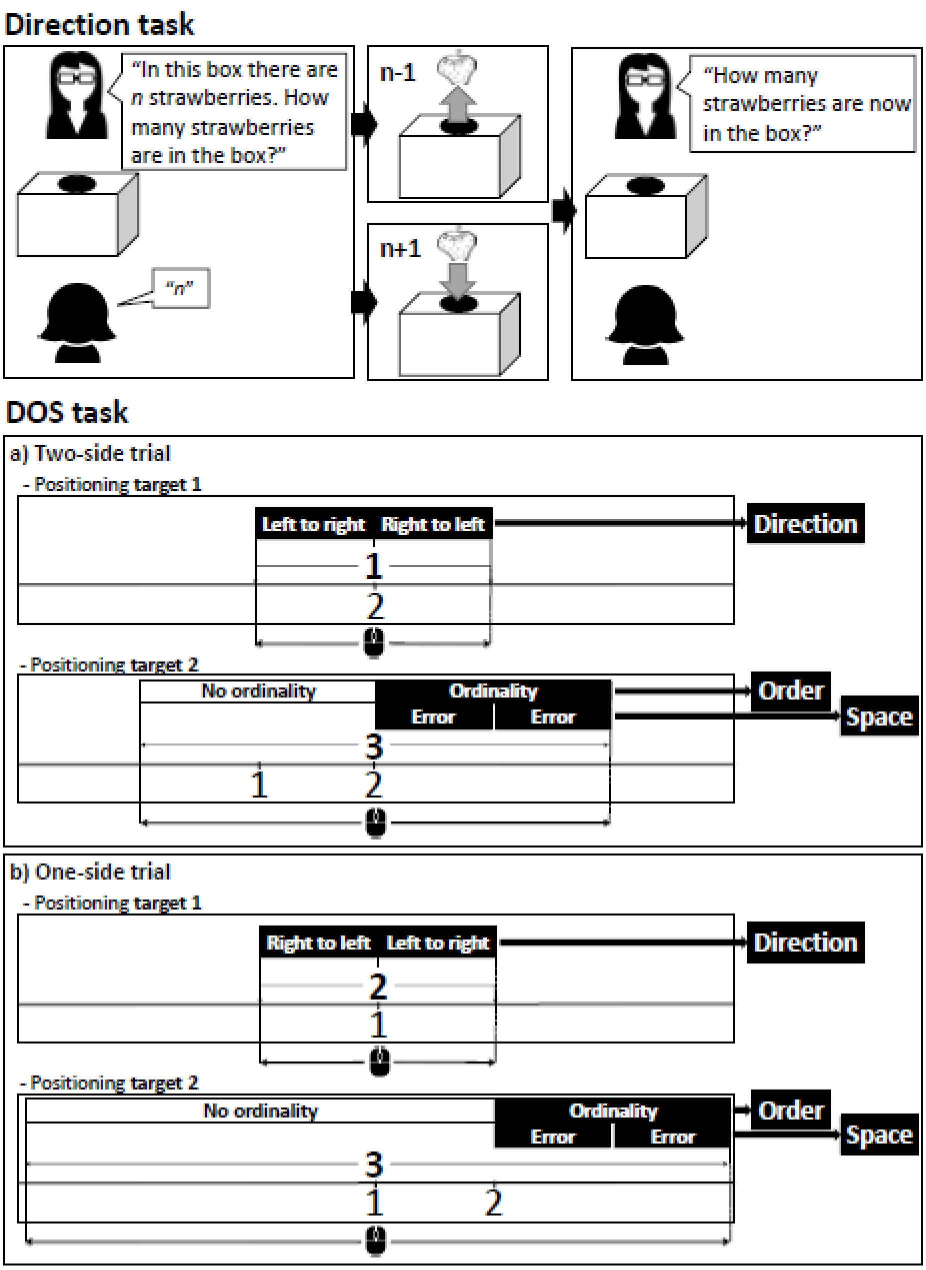

Figure 1. Top panel: Direction task. The experimenter showed a box to the child and claimed it contained a certain number of felt strawberries. Subsequently, the experimenter removed (n-1) or added (n+1) a strawberry into the box and then asked the child the number of strawberry in the box after the transformation (note: the icon representing the experimenter was obtained from www.flaticon.com). Bottom panel: DOS task. a) An example of a two-side trial for the triplet 1-2-3. The number 2 was presented in the middle of the screen just below the line (i.e., centred digit). Then, the child placed the number 1 (i.e., target 1) in one of the two designated locations, one on the left and one on the right side compared to the centred digit, thereby determining the direction of the mapping. After placing the number 1 and confirming the response, the number 3 (i.e., target 2) appeared above the centred digit and the child placed it on the line. The positioning of the number 3 determined whether the triplet of numbers was correctly ordered and the precision of the mapping. b) Example of one-side trial for the triplet 1-2-3. The structure resembled that of the two-side trials. In the

one-side trials, both targets 1 and 2 should to be placed either on the left or right side with respect to the centred number to achieve a correctly ordered triplet.

Aim the target task. This task was administered to obtain a reliable measure of children's ability in moving the mouse cursor on the computer screen. Such measure was used as a control variable when assessing the spatial mapping in the DOS task. We presented eight shooting targets (red centre with black-white-black concentric layers) one at the time in eight locations on the computer screen. The targets appeared halfway from the top of the screen and their locations varied horizontally. The cursor of the mouse could be moved on the computer screen only horizontally to facilitate the task and to mimic the same mouse restrictions implemented in the DOS task. Children were instructed to move the cursor of the mouse and click on the centre of the target. After clicking, a new shooting target appeared in a different location. For each trial, we calculated the absolute distance in pixels between the selected location and the position of the target. For each participant, we computed the median absolute distance as an index of precision in controlling the mouse.

Arabic digits comparison. Children selected the larger between two digits, which appeared respectively on the left and right side of the computer screen. There were fifteen comparisons including small pairs (i.e., 1-2, 1-3, 1-4, 2-3, 3-4), mixed pairs (i.e., 4-9, 3-6, 3-7, 2-5, 4-8) and large pairs of digits (i.e., 5-7, 5-9, 6-7, 7-8, 8-9). We presented each comparison twice with the larger digit on the right side of the screen for half of the trials. Children pressed the touchpad button corresponding to the side of the selected digit. We computed the mean proportion of correct response for each participant.

Number words comparison. The experimenter named aloud the two digits and the child said which one was the larger. The presented pairs of digits were the same used in the Arabic digit comparison task. We computed the mean proportion of correct response for each participant.

Procedure

Participants completed the tasks on two sessions usually 2-3 days apart (M=2.5, SD=1.8, range=1-8). In the first session, children completed the following tasks in this order: simple dots comparison, forward enumeration, GaN. The experimenter walked back to the classroom those children who did not meet the following criteria: 10 out of 12 correct responses in the simple dots comparison task, enumerate numbers at least up to 9 without committing mistakes, failing twice in giving one in the GaN task. Children who met these criteria also completed the aim the target task, naming and the memory tasks (i.e., Sentence Repetition and Memory for Designs) with the counterbalanced presentation of the memory tasks. In the second session, we administered the remain tasks (DOS, number words and Arabic digits comparison tasks, Direction task) in eight possible presentation orders: number words comparison, Arabic digits comparison, DOS, Direction task; Arabic digits comparison, number words comparison, DOS, Direction task; number words comparison, Arabic digits comparison, Direction task, DOS; Arabic digits comparison, number words comparison, Direction task, DOS; DOS, Direction task, number words comparison, Arabic digits comparison; DOS, Direction task, Arabic digits comparison, Arabic digits comparison; Direction task, DOS, number words comparison, Arabic digits comparison; Direction task, DOS, Arabic digits comparison, number words comparison. The computerised tasks were administered using the software E-prime 2.0 (Psychology Software Tools, 2012) on a computer laptop (monitor size: 15.6 inches; resolution 1366x768). The study was approved by the Ethics Committee for Psychology Research at the University of Padova.

Results

We ran statistical analyses using the free software R (R Core Team, 2016a) while we used the PROCESS module in SPSS (Hayes, 2013; IBM Corp., 2013) to run the mediation analyses and to estimate the 90% bootstrapped (10,000 resamples) confidence intervals for indirect effects. We pre-registered the primary statistical analyses for this study on the Open Science Framework (https://osf.io/pz8j7). Therefore, we distinguished between planned and exploratory analyses. The data and the code for the analyses can be found at https://osf.io/gczpw/?view_only=df80d2a1e5984e008eb8a11cf8860b32.

Performance in the numerical tasks

In the simple dots comparison task, children chose the numerically larger between two sets of dots whose numerical ratio was 1/2. We excluded those children with an accuracy below 10 correct responses out of 12 as the limit of a significant ($p<.05$) binomial test. In this vein, we ensure that all children in our final sample understood the meaning of more numerous, given that in the number comparison tasks children were invited to choose the larger number. Accordingly, 18 children provided ten correct responses, 17 children provided 11 correct responses, and 105 provided 12 correct responses out of 12 trials. We also administered the forward enumeration task to ensure that all children could recite the counting list up to nine given that most of the tasks, and especially the direction task, asked children to respond within this numerical range. The mean number of correctly recited number words was 23 (SD=14, range=9-100). We used the GaN task to have a robust assessment of children's cardinal knowledge, as indexed by the proportion of correct responses, and to ensure that all children were at least one-knowers. We determined the knower-level for each child by using a Bayesian classification procedure assuming the same prior probability for each knower-level (Negen, Sarnecka, & Lee, 2012). We assigned a child to a specific knower-level according to the highest peak of the posterior distribution provided by the Bayesian model. Sixty-one children were subset-knowers (18 one-knowers, 6 two-knowers, 12 three-knowers and 25 four-knowers) and 79 were CP-knowers. All children in the final sample knew the meaning of numerically more, had enough knowledge of the counting sequence to perform the direction task, and were at least one-knowers.

We calculated the proportion of correct responses in the direction task separately for n+1 and n-1 transformations. We calculated the proportion of correctly ordered triplets in the DOS task as an index of ordinality: That is, children's ability to place triplets of digits according to their ordinal relation (e.g., 1-2-3 vs 3-1-2). In the case of correct ordering in the DOS task, we calculated the corrected absolute error in placing digits controlling (residuals) for the median absolute error in pixels (M=7.6, SD=16.6) in the aim the target task, thereby ensuring that the precision in placing numbers in the DOS task was controlled for a more general ability in moving the mouse. The corrected absolute error indexes children's precision in correctly spacing numbers apart when ordinality is respected. To put it concretely, when positioning the digits 1, 2 and 3, children should place the number 3 so that the distance between 2 and 3 matches the distance between 1 and 2 (i.e., equidistance). We selected ordinality and space as these two indexes crucially assess the magnitude represented by Arabic digits whereas direction of the mapping (i.e., left-to-right or right-to-left) is arbitrary (Sella et al., 2018b). We computed the proportion of correct responses in the naming task as an additional control measure given that children independently read the digits in the Arabic digit comparison task. Finally, we calculated the proportion of correct responses in the number words comparison task and the Arabic digit comparison task as indexes of children' knowledge of exact symbolic magnitude. The detailed analyses of the performance in the direction task, the DOS task, and the number comparison tasks can be found in the supplementary materials.

In the next section, we investigated how memory and performance in the direction and DOS tasks related to number comparison skills. The descriptive statistics of the performance in the numerical and memory tasks are reported in Table 1.

Measures	M	SD	min	max
GaN task	0.78	0.24	0.29	1
Verbal memory	15.84	4.79	4	27
Memory for object location (bonus score)	12.12	10.50	0	42
Naming	0.69	0.33	0	1
Direction task (n-1)	0.55	0.31	0	1
Direction task (n+1)	0.57	0.36	0	1
DOS (ordinality)	0.66	0.21	0.17	1
DOS (corrected space)	0	0.19	-0.32	0.49
Number words comparison	0.72	0.19	0.33	1
Arabic digits comparison	0.65	0.20	0.33	1

Table 1. Descriptive statistics (N=140).

The role of directionality of the counting list and spatial mapping of numbers in mediating the relation between memory and number comparison skills

We examined how the mastering of the directional property of the counting list related to number words comparison skills after controlling for cardinal knowledge. In particular, we predicted that verbal memory specifically supports the implementation of the n+1 and n-1 transformations, which, in turn, relate to the performance in the number word comparison task. In a planned analysis, we verified whether the relation between verbal memory and number words comparison was mediated by the performance in the direction task while controlling for the accuracy in the GaN task and the memory for object location (Figure 2a-b). We found that the n+1 transformation partially mediated the relation between verbal memory and number words comparison (0.0018, 90%CI[0.0003, 0.0046]) whereas the n-1 transformation did not (0.001, 90%CI[-0.0001, 0.0037]), even though there was no significant difference between the two indirect effects (0.0008, 90%CI[-0.0019, 0.0036]).

We then examined how ordinality and precision of the spatial mapping in the DOS task related to Arabic digit comparison skills after controlling for cardinal knowledge. We predicted that memory for object location specifically supports the acquisition of the spatial mapping of digits, which, in turn, relates to the accuracy in the Arabic digits comparison task.

Therefore, in a planned analysis, we verified whether the performance in the DOS task mediated the relation between memory for object location and Arabic digits comparison while controlling for the accuracy in the GaN task and verbal memory (Figure 2d-e). We found that the ordinality of the DOS task mediated the relation between memory for object location and Arabic digits comparison (0.002, 90%CI[0.0009, 0.0035]) whereas the corrected absolute error of the DOS task did not (0.0003, 90%CI[-0.00002, 0.0012]). Accordingly, the indirect effect of ordinality was stronger than the indirect effect of corrected absolute error (0.0017, 90%CI[0.0006, 0.003]).

Despite the relation between spatial order and Arabic digits comparison, it might be conceived that children use their knowledge of the directional property of the counting list to perform the DOS task (e.g., "seven should be placed here because it comes after five and six") and to carry out the Arabic digit comparison task. To test this possibility, in an exploratory analysis, we verified whether the n+1 and n-1 transformation of the direction task, instead of the ordinality and the corrected absolute error of the DOS task, mediated the relation between memory for object location and Arabic digits comparison while controlling for the accuracy in the GaN task and verbal memory (Figure 2f). Both ordinality of the DOS task (0.0016, 90%CI[0.0006, 0.0032]) and the n-1 transformation in the direction task (0.0006, 90%CI[0.00007, 0.00154]) mediated the relation between memory for object location and Arabic digits comparison, even though the latter did not meet the standard criterion of the joint significance (Baron & Kenny, 1986). These mediation effects could be influenced by children's accuracy in naming Arabic digits, given that the experimenter did not read the digits during the Arabic digits comparison task. Nevertheless, the mediation effects remained significant also when we used the accuracy in the naming task as covariate (ordinality of the DOS task: 0.0014, 90%CI[0.0004, 0.0028]; n-1 transformation in the direction task: 0.00045, 90%CI[0.00003, 0.00139]). Finally, we verified whether ordinality and the corrected absolute error of the DOS task, instead of the n+1 and n-1 transformation of the direction task, mediated the relation between verbal memory and number words comparison while controlling for the accuracy in the GaN task and memory for object location (Figure 2c). Again, ordinality of the DOS task emerged as the only mediator of the relationship between verbal memory and number words comparison (0.004, 90%CI[0.001, 0.008]).

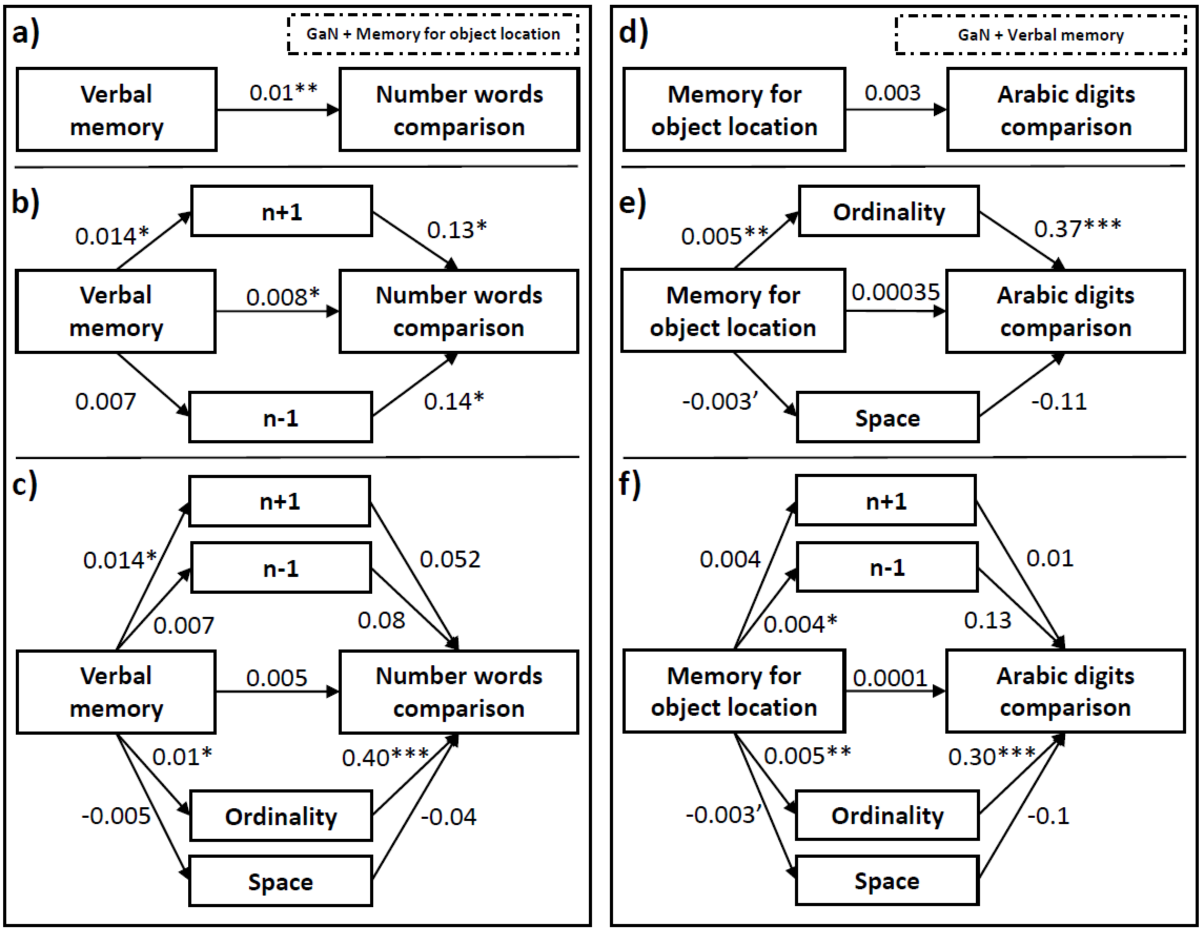

Figure 2. a) Verbal memory correlated with number words comparison while controlling for GaN accuracy and memory for object location. b) n+1 transformation in the direction task mediates the relation between verbal memory and number words comparison whereas the n-1 transformation did not. c) Only ordinality of the DOS task mediates the relation between verbal memory and number words comparison. d) Memory for object location correlated with Arabic digits comparison while controlling for the GaN accuracy and verbal memory. e) Ordinality in the DOS task mediated the relation between memory for object location and Arabic digits comparison whereas the accuracy in mapping numbers in the DOS task (Space) did not. f) Only ordinality in the DOS task mediated the relation between memory for object location and Arabic digits comparison. Variables in dotted boxes represent covariates.

Reported values are unstandardised regression coefficients. ***p<.001, **p<.01, *p<.05, `p<.10.

Discussion

In the present study, we explored different numerical skills and their relation to the ability to compare number words and Arabic digits, which we considered as the hallmark of a mature representation of the meaning of symbolic numbers. We specifically hypothesised that children achieve a full understanding of the numerical meaning of number words when they master the directional property of the counting list and use it to denote unitary operations on sets. That is, children should understand that adding one object to a set (i.e., n+1) leads to the next number word in the counting list whereas removing an item from a set (i.e., n-1) leads to the previous number word in the counting list. Indeed, the direction task required children to access the counting list at a specific position and move one step forward or backwards depending on the transformation. Such task drastically reduces the possibility that children perform the n+1 transformation as a rote behaviour, as they might do when counting forward in the GaN task. Moreover, we hypothesised that the n-1 transformation might be a crucial test given that CP-knowers already master the n+1 transformation. We also predicted that the ability to access a specific position in the counting list and perform n+1 or n-1 transformation are likely to be supported by verbal memory. The results of the present study partially confirmed our predictions: the planned analysis highlighted that verbal memory related to the n+1 transformation, which in turn related to the ability to compare number words while controlling for cardinal knowledge and visuospatial working memory. Nevertheless, the n-1 transformation contributed to performance in the number words comparison task without any mediating role.

We also hypothesised that the ability to order digits relates to Arabic digits comparison, and predicted that spatial order is supported by visuospatial memory, as indexed by memory for object location. The results confirmed our predictions. The accuracy in spatially ordering triplets of digits mediated the relation between memory for object location and the performance in the Arabic digit comparison task while controlling for cardinality knowledge and verbal memory. We also assessed whether the knowledge of the directional structure of the counting list (i.e., n+1 and n-1 transformations) could account for the mediation effect of spatial order on Arabic digits comparison. Accordingly, children could use a verbal strategy to order the triplets of digits spatially (e.g., "seven should be placed here because it comes after five and six"). This prediction was partially confirmed: both spatial order and n-1 transformation mediated the relation between memory for object location and Arabic digits comparison while controlling for cardinal knowledge and verbal memory, even though the n-1 transformation did not significantly related to Arabic digit comparison while controlling for other variables. Children's ability to name Arabic digits could have influenced the mediation effect, given that children had to read the digits during the Arabic digits comparison task. Nevertheless, the mediation effects remained significant also when we used the accuracy in the naming task as a covariate. Finally, spatial order also mediated the relation between verbal memory and number words comparison while controlling for cardinal knowledge and memory for object location. Overall, spatial order strongly related to number comparison skills while controlling for memory capacity and cardinal knowledge.

A wealth of research has emphasised the role of the mapping between non-symbolic and symbolic quantities in the construction of the symbolic representation of numbers (for a summary, Leibovich & Ansari, 2016). Children initially map small numerical quantities to the first number words and they gradually become cardinal principle knowers (Sarnecka, 2015; Wynn, 1990). CP-knowers understand that adding one item to the set corresponds to the next number words and, after few months, they also acquire the later-greater principle by creating a mapping between large external numerosities and the counting list (Le Corre, 2014; Le Corre & Carey, 2007). However, these conceptual changes in children's numerical representations only partially explain the understanding of the magnitude associated with number words and Arabic digits. Children still need to memorise the ordinal structure of the counting list and become confident in accessing it. Specifically, children who can order digits spatially (i.e., 1-2-3-4-5-6-7-8-9…) demonstrate an advanced understanding of the magnitude associated with numbers and mature knowledge of the symbolic system. The interaction between executive functions, memory in this study, and experience with the counting sequence and its visuospatial counterpart, the number line, contribute to understanding the exact quantities denoted by symbolic numbers. The cross-sectional design of the present study cannot inform on the directional relation between spatial order and magnitude knowledge. However, training studies have shown that increasing the linearity of the spatial mapping improves number comparison skills (Ramani, Siegler, & Hitti, 2012; Siegler & Ramani, 2009).

More broadly, the counting list and the number line are powerful conceptual structures to represent numerical information. An efficient memorisation and access to these structures unfold the understanding of the magnitude relation between symbols (symbol-symbol associations; Reynvoet & Sasanguie, 2016) and constitute the basis for building the first arithmetical operations. Accordingly, the speed in determining the order relation between digits has been related to arithmetic fluency in primary school children and adults (De Visscher, Szmalec, Van Der Linden, & Noël, 2015; Lyons & Ansari, 2015; Lyons, Price, Vaessen, Blomert, & Ansari, 2014; Vos, Sasanguie, Gevers, & Reynvoet, 2017). In this light, children who display poor arithmetic skills may have a deficit in memorising and efficiently accessing the structure of the number sequence. Such deficit might not be specific for numbers but extended to other non-numerical ordered sequences (e.g., letters; Morsanyi, Mahony, & Mccormack, 2016; Sasanguie, Lyons, De Smedt, & Reynvoet, 2017; Vos et al., 2017).

In summary, the acquisition of the cardinality principle is a cornerstone in the development of numerical skills. However, the status of CP-knowers does not imply a full understanding of the magnitude associated with symbolic numbers (Davidson et al., 2012; Le Corre, 2014; Sella et al., 2017). Here we showed that the ability to spatially arrange digits marks a mature knowledge of the structure of the symbolic system as indexed by the performance in the symbolic number comparison tasks. Memory, especially its visuospatial component, relates to the ability to order digits spatially. Accordingly, children gradually memorise both the shape of Arabic digits and their relative positions within the visual number line. In this regard, space can act as a powerful scaffold to build an imaginal number line (Zorzi, Priftis, & Umiltà, 2002), which constitutes a frame of reference to access the numerical magnitude of numbers (Sella et al., 2017, 2018a, 2018b).

The findings of the current study also have significant implications for education and clinical practice. We underline the importance of training young children on the ordinal component of the numerical sequence, especially learning the visuospatial arrangement of digits. For instance, Xu and LeFevre (2016) have implemented short training based on ordinality ("What number comes next/after?"), which improved basic numerical knowledge in preschool children. Similarly, the visual number line can act as a useful frame to foster children's early mathematical skills (Fischer, Moeller, Bientzle, Cress, & Nuerk, 2011; Ramani et al., 2012; Sella, Tressoldi, Lucangeli, & Zorzi, 2016; Siegler & Ramani, 2009; see also Kucian et al., 2011). Training should be designed to foster children's exploration of the number sequence, especially increasing the ability to move backwards and forward on the number line.

References

Baron, R. M., & Kenny, D. A. (1986). The Moderator-Mediator Variable Distinction in Social Psychological Research : Conceptual, Strategic, and Statistical Considerations. Journal of Personality and Social Psychology, 51(6), 1173–1182.

Barth, H. C., Starr, A., & Sullivan, J. (2009). Children's mappings of large number words to numerosities. Cognitive Development, 24(3), 248–264. https://doi.org/10.1016/j.cogdev.2009.04.001

Benoit, L., Lehalle, H., Molina, M., Tijus, C., & Jouen, F. (2013). Young children's mapping between arrays, number words, and digits. Cognition, 129(1), 95–101. https://doi.org/10.1016/j.cognition.2013.06.005

Carey, S. (2001). Cognitive Foundations of Arithmetic: Evolution and Ontogenisis. 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., 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

Cheung, P., Slusser, E., & Shusterman, A. (2016). A 6-month longitudinal study on numerical estimation in preschoolers. In Proceedings of the 38th annual conference of the cognitive science society. Austin, TX: Cognitive Science Society (pp. 2813–2818).

Cohen Kadosh, R., & Walsh, V. (2009). Numerical representation in the parietal lobes: abstract or not abstract? The Behavioral and Brain Sciences, 32(3–4), 313–328; discussion 328-73. https://doi.org/10.1017/S0140525X09990938

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

De Visscher, A., Szmalec, A., Van Der Linden, L., & Noël, M. P. (2015). Serial-order learning impairment and hypersensitivity-to-interference in dyscalculia. Cognition, 144, 38–48. https://doi.org/10.1016/j.cognition.2015.07.007

Dowker, A. (2008). Individual differences in numerical abilities in preschoolers. Developmental Science, 11(5), 650–654. https://doi.org/10.1111/j.1467-7687.2008.00713.x

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

Fischer, U., Moeller, K., Bientzle, M., Cress, U., & Nuerk, H.-C. (2011). Sensori-motor spatial training of number magnitude representation. Psychonomic Bulletin & Review, 18(1), 177–183. https://doi.org/10.3758/s13423-010-0031-3

Geary, D. C., & vanMarle, K. (2018). Growth of symbolic number knowledge accelerates after children understand cardinality. Cognition, 177(June 2017), 69–78. https://doi.org/10.1016/j.cognition.2018.04.002

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

Gunderson, E. a., Spaepen, E., & Levine, S. C. (2015). Approximate number word knowledge before the cardinal principle. Journal of Experimental Child Psychology, 130, 35–55. https://doi.org/10.1016/j.jecp.2014.09.008

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.

Hurst, M., Anderson, U., & Cordes, S. (2016). Mapping Among Number Words, Numerals, and Non-Symbolic Quantities in Preschoolers. Journal of Cognition and Development, 18(41), 41–62. https://doi.org/10.1080/15248372.2016.1228653

Jeffreys, H. (1961). Theory of probability (3rd ed.). Oxford, UK: Oxford University Press.

Jiménez Lira, C., Carver, M., Douglas, H., & LeFevre, J. A. (2017). The integration of symbolic and non-symbolic representations of exact quantity in preschool children. Cognition, 166, 382–397. https://doi.org/10.1016/j.cognition.2017.05.033

Knudsen, B., Fischer, M. H., Henning, A., & Aschersleben, G. (2015). The development of Arabic digit knowledge in 4- to 7-year-old children. Journal of Numerical Cognition, 1(1), 21–37. https://doi.org/10.5964/jnc.v1i1.4

Korkman, M., Kirk, U., & Kemp, S. (2007). NEPSY-II: A developmental neuropsychological assessment. San Antonio, TX.: The Psychological Corporation.

Kucian, K., Grond, U., Rotzer, S., Henzi, B., Schönmann, C., Plangger, F., … von Aster, M.

(2011). Mental number line training in children with developmental dyscalculia. NeuroImage, 57(3), 782–795. https://doi.org/10.1016/j.neuroimage.2011.01.070

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. (2015). Numerical Order Processing in Children: From Reversing the Distance-Effect to Predicting Arithmetic. Mind, Brain, and Education, 9(4), 207–221. https://doi.org/10.1111/mbe.12094

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

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

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

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

Psychology Software Tools, I. (2012). E-Prime 2.0. Pittsburgh.

R Core Team. (2016a). R Development Core Team. R: A Language and Environment for Statistical Computing. Retrieved from https://www.r-project.org/

R Core Team. (2016b). R Development Core Team. R: A Language and Environment for Statistical Computing. https://doi.org/http://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

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

Sarnecka, B. W. (2015). Learning to represent exact numbers. Synthese, 32(August 2015), 63–86. https://doi.org/10.1016/j.cognition.2008.05.007

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

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

Sella, F., Tressoldi, P., Lucangeli, D., & Zorzi, M. (2016). Training numerical skills with the adaptive videogame “The number Race”: A randomized controlled trial on preschoolers. Trends in Neuroscience and Education, 5(1), 20–29. https://doi.org/10.1016/j.tine.2016.02.002

Shusterman, A., Slusser, E., Halberda, J., & Odic, D. (2016). Acquisition of the Cardinal Principle Coincides with Improvement in Approximate Number System Acuity in Preschoolers. PloS One, 11(4), e0153072. https://doi.org/10.1371/journal.pone.0153072

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

Urgesi, C., Campanella, F., & Fabbro, F. (2011). NEPSY-II. Contributo alla Taratura Italiana, Giunti OS, Firenze. Firenze: Giunti O.S.

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

Wiese, H. (2003). Iconic and non-iconic stages in number development: The role of language. Trends in Cognitive Sciences, 7(9), 385–390. https://doi.org/10.1016/S1364-6613(03)00192-X

Wynn, K. (1990). Children's understanding of counting. Cognition, 36, 155–193.

Wynn, K. (1992). Children's Acquisition of the Number Words and the Counting System. Cognitive Psychology, 24(24), 220–251.

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., Priftis, K., & Umiltà, C. (2002). Neglect disrupts the mental number line. Nature, 417(May), 138–140.

Supplementary Materials

We ran statistical analyses using the free software R (R Core Team, 2016b) along with the BayesFactor package (Morey & Rouder, 2015) with default priors for Bayesian analyses. We reported Bayes factors (BF_{10}) expressing the probability of the data given H1 relative to H0 (i.e., values larger than 1 are in favour of H1 whereas values smaller than 1 are in favour of H0). When comparing regression models, we reported the Bayes factors (BF) as the ratio of BFs_{10} 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).

Performance in the direction task

We administered the direction task to examine children's understanding of the directional property of the counting list in the numerical interval ranging from 1 to 9.
The performance was analysed considering the magnitude of the starting number (small, from 2 to 4 vs large, from 5 to 8), the transformation (n+1 or n-1) and whether children were subset-knowers or CP-knowers (see Figure S1a).

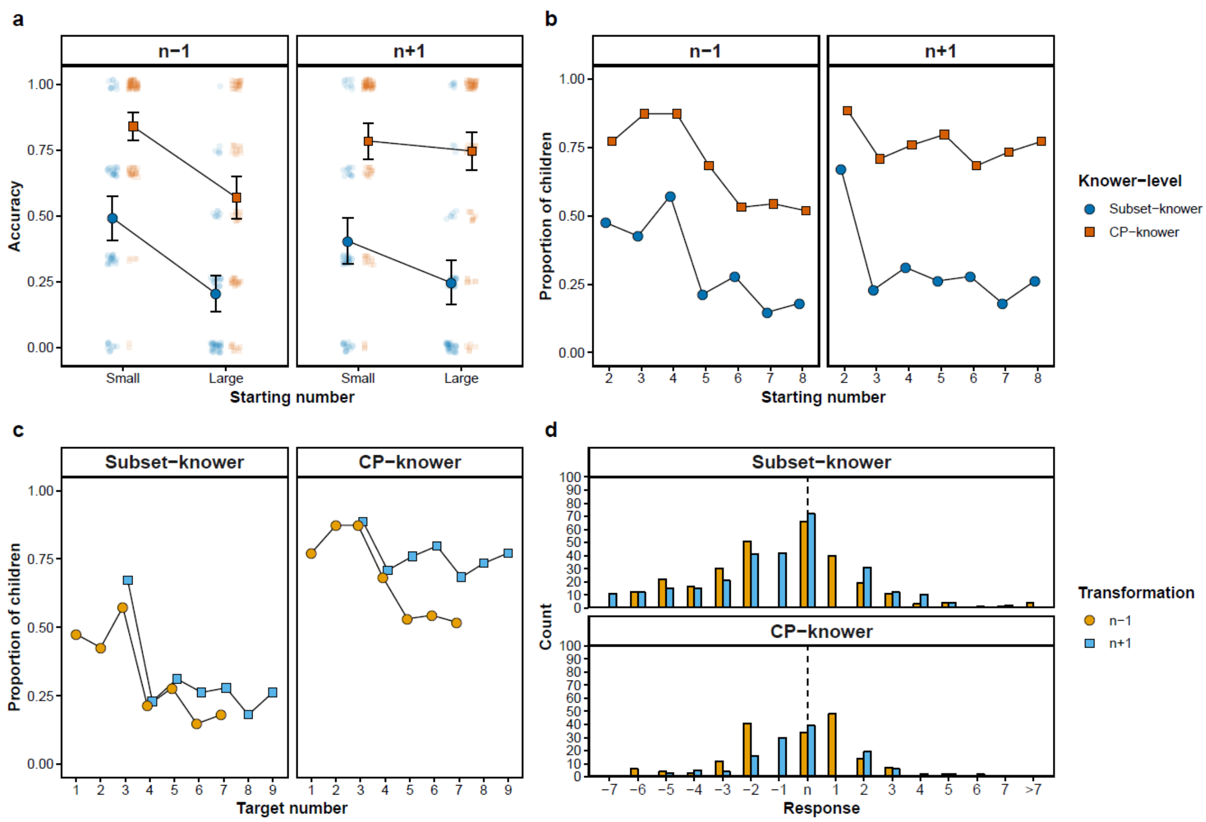

Figure S1. a) Accuracy (y-axis) in the Direction task separately for small (from 2 to 4) and large (from 5 to 8) starting numbers (x-axis) when the transformation was n+1 and n-1 separately for subset-knowers (blue dots) and CP-knowers (red squares). Transparent squares and dots represent individual values, and the error bars represent 95%CIs. b) The proportion of children who correctly performed the transformation (y-axis) in the direction task as a function of starting numbers (x-axis) when the direction was n-1 (left panel) and n+1 (right panel) separately for Subset-knowers (blue dots) and CP-knowers (red squares). c) The proportion of children who correctly performed the transformation (y-axis) in the direction task as a function of target numbers (x-axis) when the direction was n-1 (orange dots) and n+1 (blue squares) separately for Subset-knowers (left panel) and CP-knowers (right panel). d) Frequency (y-axis) of wrong responses in the direction task as a function of distance (x-axis) from the starting number (n) when the transformation was n-1 (orange bars) and n+1 (blue bars).

We ran a mixed Bayesian ANOVA on accuracy with starting number [Small, Large] and transformation [n-1, n+1] as within-subjects factor and knower-level [Subset-knower, CP-knower] as between-subjects factor. The model with the three main effects and the interaction between starting number and transformation yielded extreme evidence (BF_{10}=2.68x10^{33}) and was the most parsimonious (i.e., fewer predictors). Overall, the CP-knowers displayed better performance than subset-knowers (BF_{10}=2.97x10^{13}). The accuracy of the n+1 transformation slightly decreased when passing from small to large starting numbers (Bayesian t-test: Small n+1 vs Large n+1: BF_{10}=38, strong evidence). For the n-1 transformation, the accuracy was markedly lower for large starting numbers compared to small starting numbers (Bayesian t-test: Small n-1 vs Large n-1: $BF_{10}=7.71 \times 10^{14}$, extreme evidence).

To further examine the performance in the task across small and larger starting numbers, we plotted the accuracy in performing the n+1 and n-1 transformations as a function of starting numbers separately for Subset-knowers and CP-knowers (see Figure S1b). The performance in the n+1 transformation was higher when the starting number was two and then lower with large numbers, even though the decrease in performance was more evident for Subset-knowers than for CP-knowers. Conversely, the performance of the n-1 transformation remained relatively high for starting numbers up to 4 and then decreased for both groups. The level of knowledge of number words might explain this discrepancy between the n+1 and the n-1 transformation. In the n+1 transformation, children were asked to provide an answer that implies the knowledge of the number following the starting number whereas the n-1 transformation required the knowledge of the number that precedes the starting number. This limit in children's knowledge emerged when we plotted the performance as a function of the target number (i.e., the correct response) separately for CP-knowers and Subset-knowers (Figure S1c). Subset-knowers showed sufficiently good performance when the transformation was within three whereas their performance drastically decreased for large numbers. CP-knowers displayed proficient manipulation of numbers up to three, and then their performance remained relatively stable for the n+1 transformation and decreased for the n-1 transformation. Finally, we explored children's wrong responses in the direction task (Figure S1d). In the case of the n-1 transformation, children tended to respond by saying the starting number, the n+1 or the n-2. In the case of n+1 transformation, children frequently responded with the starting number, the n-1 and the n-2.

Finally, we plotted the individual performance in the direction task to show that the difference Subset-knowers and CP-knowers' different pattern of responses (Figure S2). There was a relevant degree of discontinuity in individual performance. In some cases, children showed some correct responses for large numbers and wrong responses for small numbers. Nevertheless, children performed overall better with small number compared to large numbers.

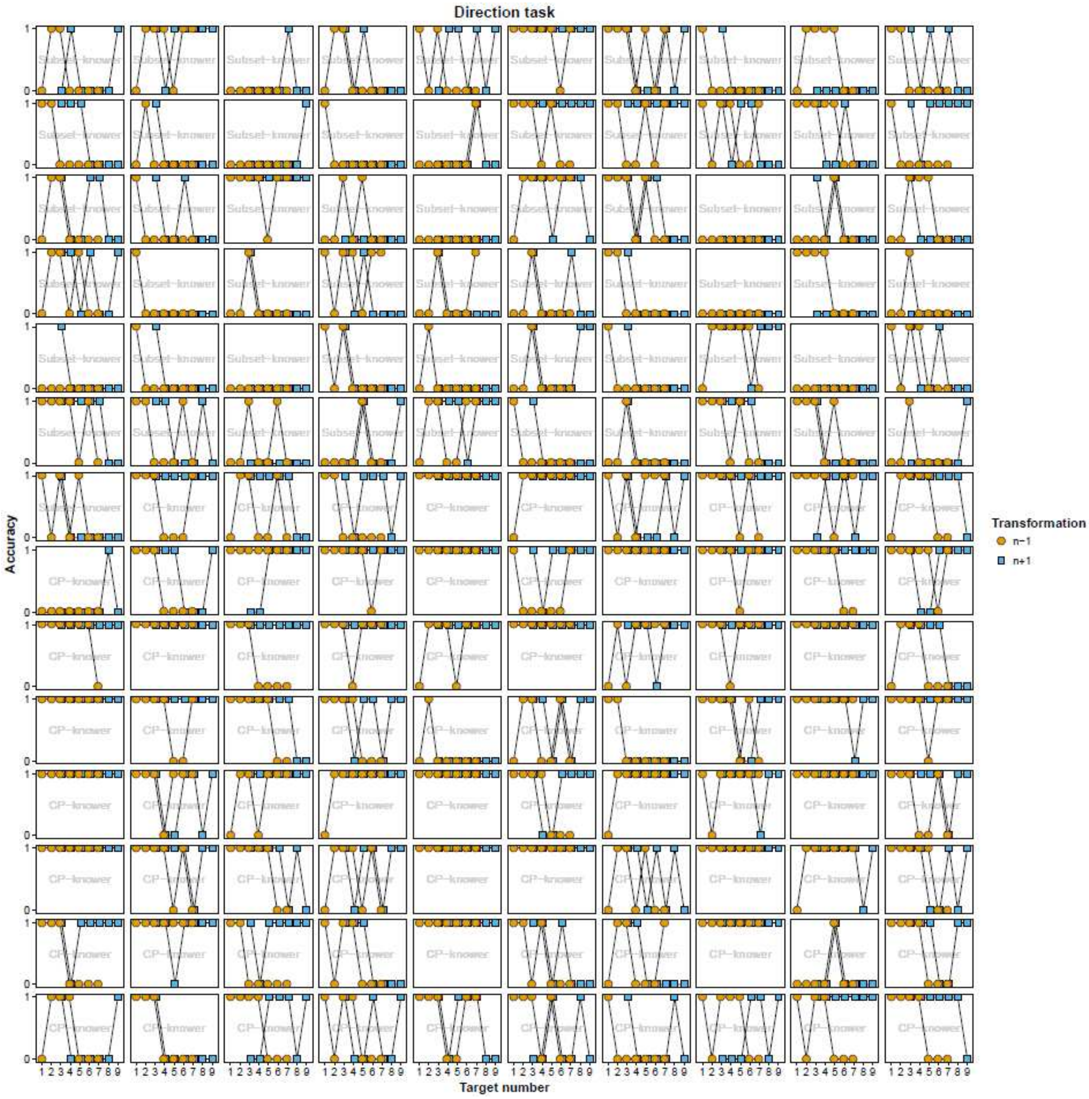

Figure S2. Individual accuracy (y-axis) in the Direction task as a function of target numbers (x-axis) when the direction was n-1 (orange dots) and n+1 (blue squares).

In summary, CP-knowers were able to solve the n+1 transformation across the entire numerical interval from 1 to 9, whereas their performance in the n-1 transformation appears to be limited to small numbers. Subset-knowers displayed an overall lower performance compared to CP-knowers. Nevertheless, subset-knowers displayed a slightly better performance with small than large numbers.

In the direction task, experimenter told the children that an opaque box contained a specific number of items before adding or removing one item. After the manipulation, children had to say the current number of items in the box. CP-knowers displayed a similar performance with small and large starting number words when the experimenter inserted an item into the box (i.e., n+1 transformation; Sarnecka & Carey, 2008). Instead, the performance for n-1 transformation decreased when passing from small to large number words. CP-knowers know that adding one item leads to the next number word in the counting list (Sarnecka & Carey, 2008) but have not fully understood that removing one item leads to the preceding number word, especially in the case of large numbers. This discrepancy between small and large number words better emerged when we explored target numbers instead of starting numbers. The target numbers identify the limit of children's numerical knowledge. For example, the same starting number three could lead a child to the unknown target number four in the case of the n+1 transformation, but the familiar target number two in case of the n-1 transformation. Approximately half of the subset-knowers completed the task for target numbers up to three and only a few of them were able to do so for larger number words. Most of CP-knowers could perform transformations up to three whereas a separation in performance between n+1 and n-1 transformations emerged for larger number words. A closer observation of individual scores (Figure S2) revealed an irregular performance. Some children could tell the numbers before and after five, failed with six, and then gave the correct answers for seven. Nevertheless, children displayed higher accuracy with small numbers as they had more experience with the initial segment of the counting list. The discrepancy for small and large target numbers suggests that children are increasing their familiarity with the counting list and, in turn, their ability to access it. Children initially memorise the order of the first numbers in the counting list and are extending their knowledge to larger numbers. The pattern of errors suggests that executive functions might support the extension of children's knowledge. For example, in the case of n-1 transformation, children tended to perform no transformation at all (i.e., repeating the starting number) or responded with either n+1 or n-2. In the n+1 transformation, again children frequently repeated the starting number or responded with n-1 or n-2. We speculate that this pattern of errors depends on executive functions, such as sustained attention, inhibition and switching. The task required children to focus their attention when the experimenter added or removed an item from the box. Missing the transformation might have prompted children to repeat the starting number. Children also had to inhibit the tendency to name the next number in the counting list as counting forward is an everyday activity for young children. They also had to switch from one transformation to the other on a trial-by-trial basis. Children who were less able to access the counting list directly possibly counted from one until they reached the target number, thereby increasing the chance to commit an error or to stop before getting to the target number. A future study that will compare the performance in a mixed design as in the current case, versus a block design (n+1 in one block and n-1 in another block), could provide further information on the contribution of executive functions to the performance observed here.

Spatial mapping of numbers in the DOS task

We administered the DOS task to obtain a detailed description of the spatial mapping of digits. A previous study has shown that CP-knowers mainly map numbers from left-to-right, their ordinal knowledge decreased with large numbers, and the precision in placing numbers is stable in the interval from 1 to 9 (Sella et al., 2018b). However, such analysis is still missing for subset-knowers. Therefore, in a planned analysis, we investigated ordinality, direction and accuracy of mapping in the DOS task as a function of triplets and cardinality knowledge (subset-knowers vs CP-knowers).

We first analysed the proportion of correctly order trials in a Bayesian mixed ANOVA with triplet [123, 456, 789] as within-subjects factor and knower-level [subset-knower, CP-knower] as between-subjects factor (Figure S3b). The model with the two main effects of triplet and knower-level yielded the largest evidence (BF_{10}=4.95x10^{17}). CP-knowers outperformed subset-knowers (BF_{10}=1.1x10^{7}) and the accuracy decreased with large triplets (Bayesian t-tests: 123 vs. 456, BF_{10}=29; 456 vs 789, BF_{10}=4764). We then calculated the proportion of trials in which children displayed a left-to-right mapping and the mean absolute error in placing number only for those trials in which children correctly ordered digits. We analysed the proportion of left-to-right mapping in a Bayesian mixed ANOVA with triplet [123, 456, 789] as within-subjects factor and knower-level [Subset-knower, CP-knower] as between-subjects factor (Figure S3a). The model with the main effect of knower-level yielded the largest evidence (BF_{10}=1.15x10^{5}). CP-knowers used more often the canonical mapping from left-to-right, whose use remained stable across triplets (BF_{10}=0.10, strong evidence towards the null hypothesis). Finally, the absolute spatial error was also analysed in a Bayesian mixed ANOVA with triplet [123, 456, 789] as within-subjects factor and knower-level [Subset-knower, CP-knower] as between-subjects factor (Figure S3c). The model with the main effect of knower-level yielded the largest evidence (BF_{10}=2.41x10^{6}). CP-knowers were more accurate in mapping digits compared with subset-knowers, and the error in mapping remained stable across triplets (BF_{10}=0.15, moderate evidence).

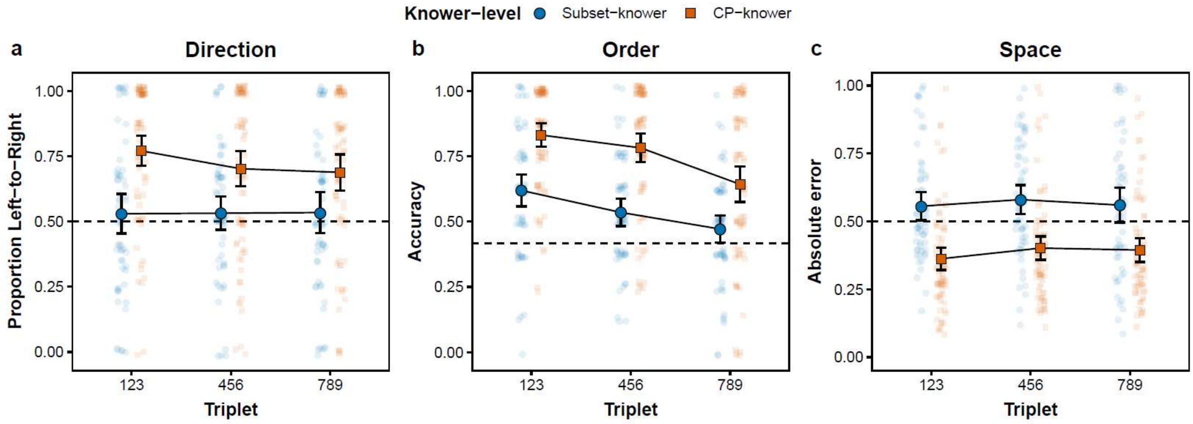

Figure S3. a) Direction: Proportion of left-to-right mapping for correctly ordered trials for each triplet. b) Order: Proportion of correctly ordered trials for each triplet. c) Space: Proportion of absolute error of spatial mapping for correctly ordered trials separately for each triplet. Mean values are presented separately for subject-knowers (blue dots) and CP-knowers (red squares). Transparent squares and dots represent individual values, and the error bars represent 95%CIs. Dashed lines represent the chance levels.

In ordered trials, the position of target 2 could vary between -1 and +1 unit from the correct position on the line (i.e., zero error). When the magnitude of target 2 was small compared to target 1 or the centred digit, a positive error value represented underestimation, and a negative value represented overestimation. Conversely, when the relative magnitude of target 2 was large, a positive value represented overestimation, and a negative value represented underestimation. Children respected the equidistance between digits when they placed the target 2 close to the correct position. Crossing under- vs over-estimation of the position of target 2 with relative magnitude (small vs large) leads to different mappings. Most of the estimates were close to zero (i.e., linear mapping) with a tendency to systematically underestimate or overestimate the position of target 2 regardless of its relative magnitude (Figure S4). Only a few children displayed a compressed or expanded mapping (for similar results, see Sella et al., 2018b).

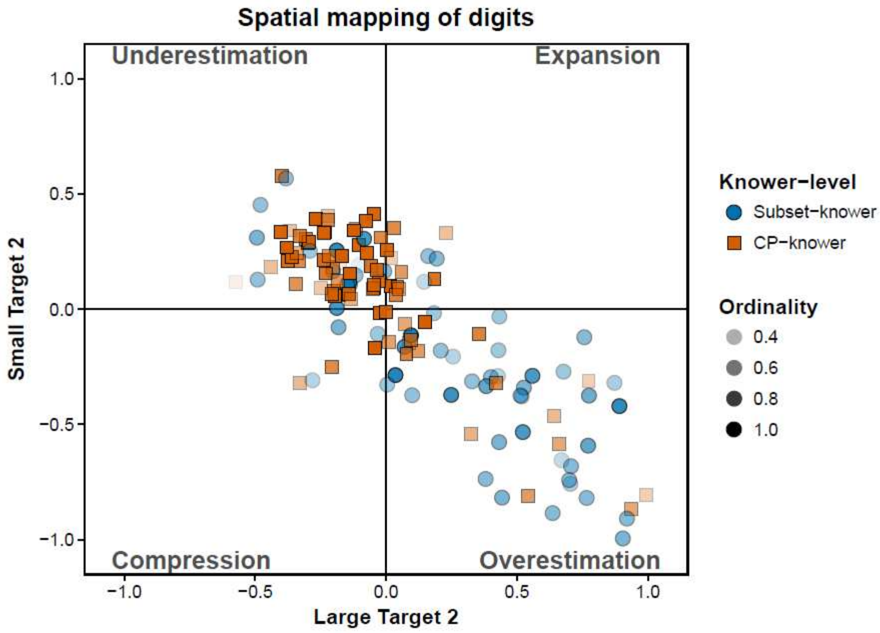

Figure S4. The Individual means of target 2 positioning for trials in which target 2 was smaller (y-axis) or larger (x-axis) compared to target 1 or the centred digit separately for Subset-knowers (blue dots) and CP-knowers (red squares). Transparency represents the proportion of correctly ordered trials.

Most studies have used the number line task to investigate young children's mapping of digits onto space (Berteletti, Lucangeli, Piazza, Dehaene, & Zorzi, 2010; Sella et al., 2017; Siegler & Ramani, 2009). Young children initially display a non-numerical or biased (logarithmic) mapping, which successively shifts to accurate and linear. For instance, subset-knowers place all the target numbers in the middle of the line or at the start and end positions, respectively 1 and 10, whereas most of CP-knowers display a linear mapping (Sella et al., 2017). However, the number line task requires children to know the proposed numerical interval and prevents a precise division between the knowledge of order and equidistance. The DOS task showed that subset-knowers could map digits on the line but only small numbers, whereas the mapping of large numbers was practically at the chance level. CP-knowers, instead, displayed an ordered mapping across the entire number line, even though their performance decreased with large numbers (Sella et al., 2018b). Children progressively increase their ability to order numbers instead of showing a clear separation in performance between small and large numbers. The direction of the mapping was mainly from left-to-right for CP-knowers whereas subset-knowers did not show a clear preference. Not surprisingly, the Italian children in our sample mapped numbers on the line from left to right according to their cultural background. Nevertheless, there was still some flexibility in the preferred mapping direction (Opfer, Thompson, & Furlong, 2010; Patro, Fischer, Nuerk, & Cress, 2016). The absolute error in mapping was constant, and the spatial mapping of digits reflected a pattern of under- and over-estimation (see Figure S4). Children ordered the three digits and placed the second target before getting to its exact position on the line. Possibly, most of the children considered the task accomplished after placing the three digits in spatial order without the necessity to respect the equidistance between digits. Overall, when taking into account ordinality, children did not display any compressed (log-like) mapping (for similar conclusions see Sella et al., 2018b).

Number comparison

We analysed the accuracy in number comparison in a Bayesian mixed ANOVA with format [Number words, Arabic digits], size of comparison [Small, Mixed, Large] as within-subjects factors and knower-level [Subset-knower, CP-knower] as between-subjects factor (see Figure S5). The model with the three main effects and with the interaction between format and size of comparison provided the highest evidence (BF_{10}=4.06x10^{32}) and was the more parsimonious. In the number words comparison task, the accuracy was higher for small and mixed comparisons than for large ones (Bayesian t-tests: Small vs Large, BF_{10}=4.78x10^{8}; Mixed vs Large, BF_{10}=8.56x10^{14}) whereas there was anecdotal evidence for a similar accuracy between small and mixed comparisons (BF_{10}=0.64). In the digit comparison task, the accuracy was higher for small and mixed comparisons compared with large ones (Bayesian t-tests: Small vs Large, BF_{10}=245; Mixed vs Large, BF_{10}=6666) whereas there was moderate evidence for a similar accuracy between Small and Mixed comparisons (BF_{10}=0.12). The interaction, which is expected based on previous research (for a review see

Cohen Kadosh & Walsh, 2009), seems to stem from a more pronounced decreasing in accuracy for large comparisons in the number words comparison tasks compared to the same difference in the Arabic digits comparison task. Overall, CP-knowers displayed a better performance compared to subset-knowers (BF_{10}=7.16x10^5).

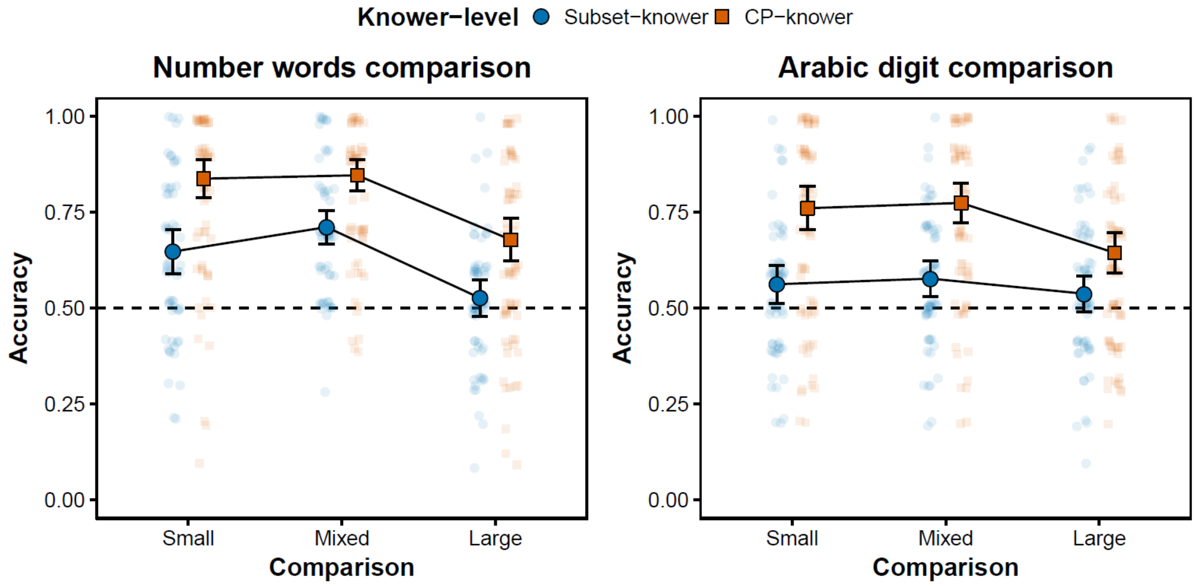

Figure S5. Accuracy in the number words comparison task (left) and the Arabic digit comparison task for small, mixed and large comparisons separately for subset-knowers and CP-knowers (red squares). Transparent squares and dots represent individual values, and the error bars represent 95%CIs. Dashed lines represent the chance levels.

Number comparison skills: Bayesian regression analyses.

In a planned analysis, we assessed the contribution of cardinal knowledge and directionality on the counting list on the ability to compare number words. We ran a regression model with the accuracy in comparing number words as dependent variable and accuracy in the GaN task, accuracy in the n+1 and the n-1 transformations of the direction task as predictors. Specifically, we compared the individual and combined contribution of n+1 and n-1 transformations after controlling for the cardinal knowledge. The model including the accuracy of the GaN task and both transformations in the direction task provided the highest evidence (BF_{10}=1.02x10^{11}, R^2=.37) compared with the other three models (i.e., GaN; GaN + Direction n+1; GaN + Direction n-1). Nevertheless, there was anecdotal evidence of its superiority compared to more parsimonious models only including accuracies of the GaN task and n+1 transformation (BF_{10}=7.05x10^{10}, R^2=.35) or GaN task and n-1 transformation (BF_{10}=3.36x10^{10}, R^2=.35).

In a planned analysis, we assessed the contribution of cardinal knowledge, ordinality, and accuracy of mapping of the DOS task on the ability to compare Arabic digits. To better assess the precision in placing numbers in the DOS task, we computed the corrected absolute error as the residuals obtained from the linear regression of the absolute error predicted by the precision in moving the mouse cursor (i.e., median absolute deviation from the aim the target task). We ran a regression model with the accuracy in comparing Arabic digits as the dependent variable and the accuracy in the GaN task, the accuracy in ordinality and the corrected absolute error of the DOS task as predictors. Specifically, we compared the individual and combined contribution of ordinality and corrected absolute error after controlling for the accuracy of the GaN task. The model including the accuracy in the GaN task and ordinality provided the highest evidence (BF_{10}=1.46x10^{10}, R^2=.34), even though there was only anecdotal evidence (BF=3.01) for its superiority compared to the model also including the corrected absolute error (BF_{10}=4.86x10^9, R^2=.34). Additionally, we also assessed the combined contribution of the n+1 and n-1 transformations in the direction task in explaining the performance in the Arabic digit comparison task above the accuracy in the GaN task, ordinality and corrected absolute error in the DOS task. The model including all predictors provided lower evidence (BF_{10}=1.36x10^9, R^2=.36) when compared to the model only including cardinality, ordinality and the corrected absolute error of the DOS task.