A refined description of initial symbolic number acquisition

Attila Krajcsi¹, Edina Fintor²

¹Department of Cognitive Psychology, Institute of Psychology, Eötvös Loránd University, Budapest, Hungary

²Department of Work and Engineering Psychology, Institute of Psychology, RWTH Aachen University, Aachen, Germany

The Give a Number (GaN) task is used to characterize the initial acquisition of symbolic numbers. Recent observations may imply that the assumptions of the original description about the development of initial symbolic numbers acquisition can be inaccurate. First, there may be children who know larger numbers than 4 and, still, they do not know all other numbers from their counting list. Second, the knowledge about these first numbers may not be all-or-nothing but it may be a more gradual knowledge. Here, a modified version of the GaN task was used to directly investigate these two properties of the development. Measuring 3- and 4-year-old children it was found that there are large-number-subset-knowers, and the knowledge about numbers is not all-or-nothing around the edge of children's known number range. These results suggest that a modified version of the GaN task would be more suitable to capture a detailed pattern of numerical development: Instead of the titration method all numbers of interest should be used, and numbers larger than 5 should also be measured. Because subset-knowers may know numbers beyond 4, these results may also challenge the Object Tracking System account of the initial symbolic number learning.

Keywords: cardinality principle; Give a Number task; subset-knowers; preschoolers' number knowledge;

Research highlight:

- There are operational subset-knowers who know larger number than 4.
- The subset-knowers' number knowledge is not all or nothing, but it is gradual.
- A modified version of the Give a number task is proposed.
- These results are challenging for the Object Tracking System account of the initial symbolic number acquisition.

1 Measuring preschoolers' initial symbolic number acquisition

Understanding how preschoolers learn the meaning of symbolic numbers is a key to understand how symbolic numbers are represented and processed in general. Children initially start to understand the meaning of number words when they are around 3 years old. Currently, acquisition of this initial symbolic number knowledge is often measured with the Give a Number (GaN) task (Wynn, 1990, 1992b). In this task, a set of similar objects are placed in front of the children, the experimenter asks for a specific number of objects, and the child should provide the appropriate set of objects. Starting from number 1, an increasing number of objects are asked until the child is not able to solve the task. Depending on whether the actual number was given correctly or not, the next asked number will be the neighboring larger or smaller number (e.g., if they fail with 5, 4 is asked, and if they succeed, 6 is asked). When a child can repeatedly give a specific number, but cannot give the next larger number, the limit of their number knowledge is determined. This procedure, termed the titration method, can determine the appropriate number knowledge because earlier measurements have demonstrated that children do not know higher numbers beyond their known number range (Wynn, 1990, 1992b). An additional criterion to categorize a number as known is that not only the child should give the correct set size for a number word, but the same number should not be given for other number words. For example, three is only known if three objects are not given for other number words, such as *four* or *five*.

In her seminal works, Wynn (1990, 1992b) convincingly described that the GaN task can be successfully used to validly specify preschoolers' number knowledge, and to specify how this knowledge develops. Initially, children cannot give any number of objects in the task, although they know the number word sequence, i.e., they can recite the one-two-three-four, etc. sequence. These children are referred to as pre-knowers or non-knowers. Once number 1 is given correctly, but not larger values, a child becomes 1-knower. Typically, a few months later, children also understand when to give 2 (2-knowers), then 3 (3-knowers), and finally 4 (4-knowers). These children are also called as subset-knowers because they can handle only a subset of their known counting list. In a next phase, children can handle all numbers that are in their counting list and can give any requested number from this list. Seemingly, at this level they understand the cardinality principle, i.e., when counting a set, the final number word is the quantity of that set, hence the term for these children, the cardinality-principle-knowers (CP-knowers). In line with this description, in the GaN task, when 5 is known, usually no larger numbers are asked, because it is assumed that these children can give all number of objects which numbers are in their counting list. It has also been demonstrated that when children become CP-knowers, they also have a general conceptual understanding what numbers are and how they could be used in simple numerical tasks (Condry & Spelke, 2008; Lipton & Spelke, 2006; Sarnecka & Carey, 2008; Wynn, 1990).

1.1 Issues with the currently used task

While the GaN task is one of the most popular task in developmental numerical cognition studies and it is the golden standard task to measure young children's initial symbolic number knowledge (Marchand & Barner, in press), there are various issues about how the task should be applied (Sella et al., 2021). The present work focuses on two specific issues: the range of the known numbers in subset-knowers, and the potential gradual nature of the subset-knowers' number knowledge.

1.1.1 Are there subset-knowers for numbers larger than 4?

Initially, it was proposed that after understanding number 4, children become CP-knowers. One main argument for this description was that no children could be observed who gave 5 but not larger values. In other words, no 5-knowers, 6-knowers, etc. could be observed (Wynn, 1990, p. 199, 1992b). Wynn (1992b) also observed that 1-2-3-knowers and the children who know larger numbers solve the task differently: While 1-2-3-knowers tend to grab the set of objects when they solve the task, other children tend to count the objects. Wynn argued that the grabbing-counting strategy reflects different understanding of numbers and this differing strategy validates the qualitative difference between children giving numbers only up to 4, and children giving numbers beyond 4. Note that this argument relies on two components: first, the GaN task performance shows a specific pattern (1-2-3-givers vs. giving all numbers in the counting list), and second, the children use different strategies (grabbing vs. counting), whereas the number knowledge and the giving strategy perfectly correlates.

However, later studies reported 5-knower, 6-knowers, or even 9-knowers (Gunderson et al., 2015; Marchand & Barner, in press; Mussolin et al., 2012; J. B. Wagner & Johnson, 2011; additionally, we also observed such children in some of our former measurements). We term these children as large-number-subset-knowers (LNS-knowers). They are subset-knowers in a sense that they can give number of objects correctly only to a specific value but not to the largest number that is in their counting list. In other words, they can be considered as subset-knowers in an operational sense. On the other hand, originally it was proposed that after knowing 4, children become CP-knowers, although they may have performance errors when handling large values. According to that proposal, and in line with the current practice of the literature, LNS-knowers should be categorized as CP-knowers even if they were not observed in the original descriptions. However, this latter interpretation is not straightforward: If 5-, 6-, or any larger number-knowers were not observed in Wynn's description and they were not validated with the counting-grabbing strategy or with other tasks, then how one should know if these children are CP-knowers or not.

At this point, it is essential to clarify the terminology and introduce new terms. We use the term operational CP-knower or operational subset-knower when children can or cannot give all the numbers that is in their counting list (independent of whether the inability of giving all numbers is caused by conceptual reasons or by performance errors). In other words, here, we do not consider the conceptual knowledge of these children, but only their performance in the GaN task. Second, we

use the terms conceptual CP-knower and conceptual subset-knower to refer to the conceptual understanding of numbers (even if conceptual CP-knowers cannot give all numbers due to performance errors and occasionally they may be operational subset-knowers). Therefore, the conceptual subset-knower and CP-knower terms consider only the conceptual understanding of the children, independent of their performance in the GaN task. Using these terms, the question we raised here is whether operational large-number-subset-knowers are conceptual subset-knowers or conceptual CP-knowers. The contradiction comes from the fact that, on one hand, the classic description found that all operational subset-knowers are conceptual subset-knowers, however, large-number-subset-knowers were not found, therefore, they were not investigated.

The studies that reported LNS-knowers categorized these children as conceptual CP-knowers and not as conceptual subset-knowers. While to our knowledge no explicit justification was given for these decisions, possibly, these works assume that in line with Wynn's argument, children who know larger numbers than 4 are CP-knowers who make performance errors. Note again that operational LNS-knowers were not observed and validated in Wynn's description, therefore, one cannot be sure if these children are conceptual CP-knowers.

Regarding a related methodological issue, in Wynn (1992b), only a single 4-knower was found and she was categorized as CP-knower because e.g., she was a counter. Therefore, the classic description proposes that 4-knowers are conceptual CP-knowers (or at least the conceptual or validational viewpoints should be considered). Still, the literature usually categorizes 4-knowers as conceptual subset-knowers, while at the same time these works refer to Wynn's papers to justify the evaluation procedure. To our knowledge this discrepancy is not explained in the literature either, but one possible reason can be that because it was originally observed that operational subset-knowers are usually conceptual subset-knowers as well, and because 4-knowers are observed relatively frequently, these children are categorized as conceptual subset-knowers.

Overall, there are contradictions between Wynn's description and argument, and later observations and practices. In Wynn's description, LNS-knowers are assumed to be nonexistent, however, LNS-knowers later have been observed. Wynn argued that validational viewpoint, e.g., the counting-grabbing strategy should be primary and, accordingly, 4-knowers can be CP-knowers, however, later practice relies only on the given numbers. Finally, Wynn found that operational subset-knowers are always dominantly grabbers, and children who give all numbers in their counting list are always dominantly counters. However, it is not known that if LNS-knowers exist what their conceptual knowledge is.

Although relying on the validational tasks in these decisions seems to be reasonable, there are issues with this possibility. First, the whole argument that suggests that 4-knowers are conceptual CP-knowers is based on a single child (footnote 2 in Wynn, 1992b). Based on single child it cannot be known if her case was typical or not. Second, in the literature, there is a debate about whether the validational tasks can differentiate subset-knowers and CP-knowers appropriately. While several studies argue that CP-knowers show a qualitatively different knowledge of the number properties

compared to subset-knowers (Condry & Spelke, 2008; Lipton & Spelke, 2006; Sarnecka & Carey, 2008; Wynn, 1990, 1992b), some others argue that CP-knowers perform more poorly than it would be expected if a child had an appropriate conceptual understanding of numbers. For example, it has been demonstrated that some of the CP-knowers cannot estimate the numerosity of a visual set (Le Corre & Carey, 2007), they cannot map symbolic numbers into a spatial number line (Sella et al., 2017), they cannot compare symbolic numbers (Davidson et al., 2012; Le Corre, 2014; Sella et al., 2017), or they do not know how a set is changed when an item is added or removed (Davidson et al., 2012; Sella & Lucangeli, 2020). Overall, while validation viewpoints may help to determine if children are subset-knowers or CP-knowers, the results are not straightforward, and there is no consensus in the literature which task could be optimal for appropriate validation.

Similarly, even if the pattern of the given numbers in the GaN task in itself is informative, there are related issues. It is possible that operational LNS-knowers were not observed in the classic description because of methodological constraints. First, the classic description included small samples, and there is a chance that some rare categories could not be observed only because of random sampling. For example, the relevant third experiment in Wynn (1990) included only 18 children, and Wynn (1992b) included 20 children, while only a single 4-knowers was reported in the two papers. Second, only a limited number range was used in the GaN task. In the third experiment of Wynn (1990), 1, 2, 3, 5, and 6 were used, therefore, all LNS-knowers who know at least 6 could look like CP-knowers (note also that 4-knowers could be categorized as 3-knowers, because 4 was not asked). Similarly, in Wynn (1992b), only the numbers between 1 and 5 were used, so LNS-knowers were impossible to observe. This limitation applies to most later studies as well: Either no larger numbers than 6 are asked, or even if they are asked, the relevant results are not reported. Finally, while, in some studies, LNS-knowers were reported (Gunderson et al., 2015; Marchand & Barner, in press; Mussolin et al., 2012; J. B. Wagner & Johnson, 2011), these incidental reports usually did not present sufficient relevant details about these rarely reported patterns, and the limited reliability of the GaN task (Marchand & Barner, in press) may question how reliable those reports can be. Overall, in the classic description, the samples were too small and the applied numbers were limited to have appropriate chance to observe LNS-knowers, and later accidental reports did not provide sufficient details about potential LNS-knowers. Let us highlight here again that while validation tasks could be essential, the operational viewpoint is also essential in itself because the observed seeming lack of LNS-knowers was one of the main arguments why children can be categorized into subset-knowers and CP-knowers in a conceptual sense – however, the methodological constraints mentioned here seriously limited what results could be observed.

To summarize, while it was originally proposed that there are no operational subset-knowers beyond 4, LNS-knowers have been reported in the literature. According to the current practice in the literature these operational LNS-knowers are categorized as conceptual CP-knowers. However, it is not straightforward whether this categorization is justified because of the limitations of Wynn's description and because of the contradictions between Wynn's description and later practices. To clarify the status of the LNS-knowers, both operational and conceptual viewpoints should be

considered. As a first step, the present work focuses on the operational viewpoint and investigates whether LNS-knowers can be observed, the detailed performance pattern behind their knowledge, and how common they could be. Note that while the validation or conceptual question is equally essential, here, we focus only on the operational question, and currently leave it to other works to consider the issues of validation.

1.1.2 Is the subset-knowers' number knowledge all-or-nothing?

In the seminal description, Wynn (Wynn, 1990, 1992b) found that measured with the GaN task subset-knowers know the numbers up to a specific value, but they do not know larger numbers. Following this finding, the procedure that is used in the literature assumes that these children's knowledge is all-or-nothing: They either know the number consistently or they do not know the number. However, some findings suggest that this knowledge can be more gradual around the edge of the number range that children know reliably.

It was found in several studies that the GaN task responses show that children have partial knowledge beyond their knower level: A subset-knower may partially know the value of the next two "unknown" values in their counting list (Barner & Bachrach, 2010; Gunderson et al., 2015; K. Wagner et al., 2019). O'Rear, McNeil and Kirkland (2020) also found that children learn the next number faster if they have partial knowledge of that number (i.e., the child could give appropriate number for a number word, e.g., three objects for "three", but the same number was erroneously given for other number words, as well, e.g., three objects for "four") compared to children with no knowledge of that number (the appropriate number cannot be given correctly at all). These results mean that the number knowledge may not be all-or-nothing, but beyond the children's measured knower level, there may be partial knowledge about the next values.

As a related issue, Sarnecka and Gelman (2004) reported 4 children out of a sample of 54 children who had "gap" in their knowledge: Some children seemed to know a number range, although some of the values were not known within that range. In line with this finding, we also observed similar gaps in some of our former measurements, e.g., one of the preschooler participants could not give number 3, but could give number 4 consistently. Note that it is impossible to observe such gaps with the titration method because the titration method stops at the first "unknown" number. This also means that the titration method may underestimate the knowledge of the children when there is a gap in their number known number range. One critical question is why such gaps can appear in the children's performance. On one hand, these gaps can be simply random performance errors. On the other hand, these gaps could be related to a more gradual knowledge of numbers, where gradual knowledge measured with random noise will make correct responses randomly be above or below the threshold criterion of knowing that number.

Overall, previous findings suggest that the number knowledge of the children may not be all-or-nothing, but it can be more gradual, and number knowledge with gradual knowledge, or with gaps, or with performance errors may not be captured appropriately by the titration method.

1.2 Aims of the studies

The first aim of these studies is to verify if LNS-knowers can be observed. The seminal description of the symbolic number development measured with the GaN task found that LNS-knowers cannot be observed, however, later accidental observations reported opposing results. Therefore, as a first step to clarify the observable patterns, the possible presence of the LNS-knowers should directly be investigated. Note again that it is not the aim of these studies to investigate the conceptual knowledge of the possible operational LNS-knowers. The second aim of these studies is to investigate whether the knowledge of subset-knowers is all-or-nothing or, alternatively, a more gradual transition between known and unknown numbers can be observed. The original description of knower levels suggested that the knowledge of subset-knowers is all-or-nothing, while later findings may question this assumptions. Because the literature mostly relies on the classic description of the GaN task despite the conflicting findings or the methodological limitations of that description, it is essential to clarify these phenomena. To draw a clearer picture, these studies use a modified version of the GaN task.

2 Study 1

2.1 Methods

Three- and four-years-old preschoolers performed the GaN task where numbers from 1 to 9 were asked four times. Asking numbers up to 9 enables us to investigate whether there are 5-, 6-, 7- or 8-knowers, while former studies usually measuring number knowledge up to 5 could not differentiate operational LNS-knowers from CP-knowers. Additionally, asking all numbers in the given range enables us to discover partial knowledge of a number that could not be seen with the titration method. Asking all numbers four times could result in a more reliable data compared to the former use of the GaN task.

Previously, we ran a similar preparatory study which we describe in the Supplementary material. The findings of that study were basically the same as in the present study.

The study was approved by the ethics committee of the Cognitive Psychology Department, Eötvös Loránd University, “Investigating numerical development in preschoolers”.

2.1.1 Participants

Seventy-two preschoolers participated in the study, 28 girls and 44 boys, with a mean age of 49.1 months, SD of 5.7 months, ranging between 37 and 60 months. Nineteen Hungarian preschools were involved in the measurement, 8 of them in the capital, and 11 of them in a country town, most of them receiving children of middle-class families, and three of them having children from above average or below average SES families.

2.1.2 Tasks and procedure

Give a number task. The children were presented with a pile of sponge balls (approximately 30 balls, with a diameter of around 2 cm), and the experimenter asked the children to give a specific number of balls. “In this task, I want to see if you can select these balls. Could you give me x balls? Just place them in front of me.” At the end of the trial, it was asked “Is this x balls?” If the child replied with “no”, then the set was asked again: “Give me x balls then.” No feedback was given after a trial. Instead of the titration method, all numbers between 1 and 9 were asked, therefore, it was possible to investigate if children have gradual knowledge of numbers. Also, this wider number range enabled us to potentially observe LNS-knowers. The values were presented in the following order: 1, 3, 2, 6, 4, 5, 7, 9, 8. The series of the 9 numbers were asked 4 times, resulting in 36 trials. This order was used to overcome the predictive nature of an increasing series. The given numbers were recorded.

Counting list knowledge. The child was asked to continue the counting list: “Could you continue this series? I start it, then you continue! One, two, three...” The task was asked twice. The highest number while the series was correct was recorded.

The two tasks were measured in a single session. First the Give a number task, then the Counting list knowledge task was asked. Measurement time of a child was around 15-20 minutes. Data were collected by ten experimenters, all of them blind to the question of the study.

2.1.3 Analysis

Calculating number knowledge in the Give a number task. Average performance for all numbers was calculated for all children. According to a first criterion, a number was considered to be known if the performance was higher than 50% (75% or 100%), and a number was categorized as not known for worse performance (50%, 25%, or 0%). This criterion is in line with the original, titration version, where it was either tested whether “child succeeded at reliably” after the numbers at the edge of the participants’ number knowledge were asked three times (Wynn, 1992b), or a number was evaluated as not known when the number was not given at least two out of three times (Le Corre & Carey, 2007) and when the previous number was given at least two out of three times (Cheung et al., 2016). As an additional criterion, a number was considered as known only if the same set size was not given for other numbers more than twice. For example, frequently, the largest correctly used number is also given for the next number word, e.g., when three is given correctly, some children give three for the number word “four”, too. In that case, the number three is not evaluated as known, because the set of three is not entirely understood.

After specifying the known numbers, the number knowledge was calculated with two methods. (1) Unlike in the titration version of the GaN task, the present version asked all numbers in the 1-9 range, independent of whether smaller values were given correctly or not (see similar methods in Marchand & Barner, in press; Mussolin et al., 2012; Sarnecka & Gelman, 2004; Sella & Lucangeli, 2020; J. B. Wagner & Johnson, 2011). To imitate the classic titration method, number knowledge

was specified as the largest known number, after which the first unknown number comes. In other words, number knowledge is the smallest unknown number minus 1. For example, if the smallest unknown number is 4, then the number knowledge of that child is 3. This calculation method follows the titration method, because in the titration method when an unknown number is reached, no further numbers are asked. (2) To overcome the potential limitation of the titration method that cannot find known numbers after a number is unknown (i.e., ignores possible number knowledge after a gap), an alternative index was also calculated. As a simple alternative index, the largest known number was considered as the number knowledge, even if there were smaller unknown numbers. For example, if 1, 2, and 4 were known, but 3 was unknown, then the child was categorized as a 4-knower. This calculation method was also used by Wagner and Johnson (2011), although the use of this method was not justified. The two methods would show different results only if there is a gap in the child's number knowledge, otherwise the two methods should show the same result.

Calculating counting list knowledge. Counting list knowledge was calculated as the mean of the highest numbers recited correctly from the counting list in the two tasks. Mean was considered as a compromise between minimum reliable performance (minimum of the two values, i.e., that number was reached consistently) and the maximum peak performance (maximum of the two values, i.e., a value that is not always reached).

2.2 Results and discussion

The raw data can be found at <https://osf.io/z28hd/>.

In the counting list knowledge task, 48 (67.6%) children could recite the counting list at least up to 9. In one other case, the child could not solve the task. There were only three preschoolers whose counting list knowledge was smaller than their number knowledge measured with the largest known number calculation method (they were 3-, and 6-knowers). For the remaining 20 children the counting list knowledge was higher than their number knowledge, but lower than 9. Overall, for the majority of the children, counting list knowledge did not constrain their number knowledge.

When, at the end of the trials, it was asked whether the given set was the appropriate amount, the response was corrected on average 8,5% of the trials. 45.8% of the preschoolers never corrected their answer.

2.2.1 LNS-knowers

In the present sample, LNS-knowers could be observed (Figures 1C and 2) and all possible LNS-knowers, i.e., 5-, 6-, 7-, and 8-knowers, were present, independent of the number knowledge calculation method (Figure 2). In the present sample, there were one 5-knower, one 6-knower, two 7-knowers, and two 8-knowers, all together six LNS-knowers. Except one preschooler, all of these children could recite the counting list to higher number than their number knowledge, therefore, the limited counting list knowledge did not constrain their number knowledge. Finding six LNS-

knowers suggest that this may not be a typical category in knower level, e.g. compared to the nine 4-knowers in the same sample. However, these relatively rare categories may fluctuate even in a relatively large sample: In the preparatory study, where only 5-knowers were measured as LNS-knowers, five 5-knowers were found compared to a single 4-knower (see the Supplementary material for details). Overall, while LNS-knowers are not common, the response pattern exists. A more detailed analysis of the responses is given in the next subsection.

This result is in line with previous reports in the literature (Gunderson et al., 2015; Marchand & Barner, in press; Mussolin et al., 2012; J. B. Wagner & Johnson, 2011), and it is in contrast with the initial observation of Wynn (1990, 1992b) where no LNS-knowers were found. This latter point is essential because the lack of the LNS-knowers was one of the main reasons why it is assumed that children become CP-knowers after giving 4 correctly in the GaN task.

Note that the present data are coherent with the data of Wynn (1990, 1992b) in a sense that if we followed the methods of those works, we would have the same results as in those descriptions: If numbers are asked only up to 5 or 6, and if the sample is relatively small, then there is a chance that 4-, or 5-knowers could not be observed, and, 6-, 7-, and 8-knowers would be impossible to find. In a parsimonious interpretation, it is sufficient to assume that the difference of the methods can explain why the present finding differs from Wynn's finding.

This result highlights that larger numbers should also be measured in the GaN task to have a more comprehensive picture about preschoolers' symbolic number understanding. We discuss the issue of these children's conceptual knowledge and the implications for developmental theories in the General discussion.

2.2.2 Partial knowledge of numbers

To see if subset-knowers' knowledge about numbers is all-or-nothing or whether it can be partial, individual data were investigated. Additionally, more detailed individual data may shed light on possible response pattern differences between SNS-knowers and LNS-knowers. In Figure 1, individual data are displayed, each row displaying a single participant's data. Children are ordered according to their overall performance (B) – this ordering is purely data-driven, no categorization is involved in it, therefore, it is neutral from the viewpoint of the two number knowledge categorization methods (C).

Figure 1. Individual performance in the Give a number task. A row includes the data of a single child. In the columns, (A) mean correct responses for the numbers from 1 to 9 (green showing 100% and white showing 0% performance), (B) mean performance for the whole task (mean of the columns in (A)), and (C) number knowledge calculated as the smallest unknown number – 1 and

calculated as the largest known number (0 meaning pre-knowers) (darker color reflecting larger number knowledge).

The leftmost part of Figure 1 displays the mean performance for a single value (the value could be 0%, 25%, 50%, 75%, or 100%, because all numbers were asked four times). First of all, it was found that in line with former suggestions and the classic titration-like number knowledge calculation index, larger numbers are usually only known when smaller numbers are already learned (this can be seen in Figure 1 as numbers on the left are green, and numbers on the right are white, and no reversed pattern or green-white-green pattern can be seen)¹.

However, it is also noticeable that there are several children (usually children with mid-level average performance in our sample) who show an unstable performance for several values (lighter green cells for several values): These children firmly know the smallest values, then show an uncertain performance for some larger values (but still above some very low chance level), and finally they show a complete lack of understanding of the highest numbers in the present task. This pattern shows that children's knowledge is gradual in a specific range of the values: While knowledge for a relatively small number can be firm, and large numbers are hardly known at all, there is quite a large range of numbers where the performance is somewhere in between. Importantly, it would be hard to specify where the boundaries of these known–half-known–unknown number ranges lie. In the half-known range, seemingly random fluctuations can be observed, which may explain why gaps can be observed: With this random fluctuation in the half-known range, some numbers will be categorized as known number, and some of them unknown, where known number can follow unknown number. This aspect also suggests that the titration method is not necessarily appropriate to measure children's number knowledge.

It is also observable that range of the uncertain numbers is larger for children knowing more numbers than for children knowing less numbers. This pattern may explain why reliability for larger numbers is smaller than for smaller numbers (Marchand & Barner, in press). For an uncertain number in the GaN task, children sometimes give the correct response, and sometimes they give the incorrect response. This unstable response leads to lower reliability, since the response for the same number may change across trials and, consequently, across sessions. Because larger numbers show larger uncertainty, this can be observed as lower reliability for those numbers. Therefore, the observed gradual knowledge of some values and the larger uncertainty of larger numbers can explain why the reliability of larger numbers can be lower.

It is also worth to note that looking at only the response patterns of these children, no characteristic difference can be observed between SNS-knowers and LNS-knowers. While it is possible that SNS-knowers fail on larger numbers because they lack appropriate knowledge, and LNS-knowers fail on larger numbers because of performance errors (i.e., there could be difference in their conceptual

¹In some cases, the number knowledge (Figure 1C) is smaller than the number of known numbers (Figure 1A). This can happen when a number was given correctly, but typically the same number was given for the next larger number word as well, e.g., although 3 balls were given for the number word “three”, but 3 balls were given for the number word “four”, too, and therefore, the number 3 is categorized as unknown.

knowledge), still, the pattern of the mean correct responses does not show qualitative differences between these two groups. Note, however, that these two patterns do not have to be qualitatively different even if there are different generators for that pattern. For example, SNS-knowers may give partial response for unknown numbers based on the Approximate Number System (ANS) which may provide responses with a pattern of normal distribution (see more details in the General discussion), while LNS-knowers' performance error increases with the value of the number where more steps in the counting procedure increases the probability of making an error, resulting in a binomially distributed response pattern. In these cases, the ANS-based normal distribution and the performance error-based binomial distribution can be indistinguishable. To summarize, the response pattern does not provide any evidence that the SNS- and LNS-knowers behave qualitatively differently. Validation studies can investigate whether the two groups have qualitative difference in their conceptual understanding of numbers despite the similar response patterns.

To investigate how differently the titration method and our simple alternative number knowledge index (i.e., which considers the largest number) measure number knowledge, a scatterplot of the two indexes is shown in Figure 2. The figure displays the number of children with a specific number knowledge level measured with the classic calculation methods (x-axis) and measured with the new largest known number index (y-axis). (The same information can be seen in numerical format in Figure 1C.) Difference between the two indexes means gap in the children's number knowledge, i.e., children who are not on the $y = x$ diagonal have different number knowledge measured with the two indexes. Results show that most of the children were categorized to the same number knowledge level according to the two calculation methods, as seen in the points lying across the $y = x$ line. However, 9 preschoolers (12.5%) were categorized differently by the two methods – these children showed a gap in their knowledge.

Figure 2. Number of children with a specific number knowledge calculated with the titration-like method (smallest unknown number – 1) (x-axis) vs. calculated with the largest known number method (y-axis). Zero number knowledge means that a child is a pre-knower. Largest size denotes 19 children, smallest size denotes 1 children. Children with the same number knowledge calculated by both methods lie along the $y = x$ diagonal.

Finally, note again that a preparatory study with similar methods has also been run, which has found basically the same findings. See its details in the Supplementary material.

Overall, these results are in line with former observations. First, after the known numbers, half-known numbers can often be observed, which explains why the next 1 or 2 values beyond the classic number level was found to be partially known (Barner & Bachrach, 2010; Gunderson et al., 2015; O’Rear et al., 2020; K. Wagner et al., 2019). Second, gaps were observed as a consequence of fluctuations in the half-known number range, which explains the observed gaps in children’s number knowledge (Sarnecka & Gelman, 2004). Finally, the observed gradual knowledge is in line with the lower reliability of larger numbers of the GaN task (Marchand & Barner, in press).

This pattern also reveals why the methods that are currently used in the literature cannot capture these properties. First, numbers are categorized as known and unknown numbers instead of detecting gradual changes in number knowledge. Second, the titration method will stop at the first number categorized as unknown, and can underestimate the children’s number knowledge and it can not be able to measure the additional knowledge beyond that first unknown number.

Based only on the pattern seen here, it is hard to determine which index could capture number knowledge more appropriately. Is not the largest known number index too liberal in a sense that unstable knowledge is also included? Should the partial knowledge beyond the largest known number be considered? Is not the classic titration method related index show the reliable knowledge

(Marchand & Barner, in press)? While based on the present results it is impossible to determine the ideal solution, some considerations are discussed in the General discussion.

3 Study 2

Based on Study 1, it is hard to specify why some numbers are half-known. Is it an emerging knowledge that is not complete yet? Is it an unstable knowledge where performance error is more probable? Recently, it has been demonstrated that the follow-up question at the end of the GaN task trial influences the performance: Children show higher number knowledge if they are asked to recount the given set after their response compared to when they are only asked if the given set was the same as the required set or when no follow-up question is asked (Krajcsi, 2021). However, it is not yet known whether the performance is better because recounting may decrease performance errors (i.e., children understand the task conceptually, although they make error because they cannot execute the counting procedure flawlessly, as assumed in Wynn (1990, 1992b)) or because recounting provides a training for the children (as in Mix et al. (2012); see related examples in Krajcsi (2021)). If recounting decreases the performance error, then the result of Study 1 in which it was only asked whether the appropriate set was given (or the result of the study described in the Supplementary material which had no follow-up question) may be biased: In these studies, some conceptual CP-knower children may show performance errors and they may look like LNS-knowers, although they are only operational LNS-knowers.

To investigate whether the lack of a recounting instruction may cause the appearance of operational LNS-knowers, a new dataset was collected where, at the end of the GaN trials, children were asked to recount the given set to be sure that they performed the task correctly (Wynn, 1990). If the LNS-knowers that were observed in the previous study are conceptual CP-knowers who could not perform the task appropriately, and if the recounting task version decreases the performance error, then this recounting task version can make the LNS-knowers disappear.

3.1 Methods

This study used the same methods as in the previous study except that, at the end of each trials, the experimenter asked the children to recount the given set. If the result was different from the required amount, the required amount was asked for again. With this GaN task version children could potentially detect and correct their errors.

3.1.1 Participants

Seventy-one preschoolers participated in the study, 42 girls and 29 boys, with a mean age of 49.5 months, SD of 6.4 months, ranging between 37 and 59 months. The same preschools were involved as in the first study.

3.2 Results and discussion

Like in Study 1, for the majority of the children, counting list knowledge did not constrain the number knowledge. In the counting list knowledge task, 49 (70.0%) children could recite the counting list at least up to 9. There were three children whose counting list knowledge was smaller than their number knowledge measured with the largest known number calculation method (they were 3-, 8-, and 9-knowers). For the remaining 19 children the counting list knowledge was higher than their number knowledge, but lower than 9.

As in the previous study, LNS-knowers could be observed (Figure 3C and D, and Figure 4). Calculated with the titration-like method, there were four 5-knowers, five 6-knowers, two 7-knowers, and one 8-knower (17% of the sample). Similarly, if the calculation relies on the largest known number, LNS-knowers can be observed as well.

Similar to the first study, the gradual nature of number knowledge was also observable: performance decreases gradually for larger numbers instead of a sudden performance drop.

As in the first study, depending on the number knowledge calculation methods, some children may be categorized differently: 12.7% of the preschoolers (9 of 71) were categorized differently (Figure 4).

Figure 3 Individual performance in the Give a number task. See details in Figure 1.

Figure 4 Number of children with a specific number knowledge calculated with the titration-like method (x-axis) vs. calculated with the largest known number method (y-axis). Largest size denotes 23 children, smallest size denotes 1 children. See more details in Figure 2.

Study 2 demonstrated that even if the recounting instruction at the end of the GaN trials potentially decreases the number of performance errors, LNS-knowers are still observable. Again, these results are in line with previous reports in the literature observing LNS-knowers, and they are in contrast with the initial reports of Wynn (1990, 1992b).

Note that it is possible that the current procedure that the literature uses to decrease performance error may not eliminate all performance errors, and the present results may still be biased. This would also mean that the reported results of other studies using this recounting version of the GaN task are also biased. Future studies may investigate this possibility. On the contrary, it is also possible that the recounting version does not prevent performance error, but provides training for the GaN task (Krajcsi, 2021). In this case, the methods of Study 1 would be more appropriate. Importantly, the results of Study 1 and 2 show the same pattern.

4 General discussion

The present work investigated two properties of the initial symbolic number acquisition in preschoolers, measured with the GaN task: (1) Are there operational LNS-knowers, i.e., children who may know larger numbers than 4, but not all numbers in their counting list, and (2) is children's number knowledge gradual (i.e., not all-or-nothing) that is not visible for the GaN

method applied in its current form. None of these phenomena could be investigated with the conventional version of the GaN task, therefore, a wider range of number (between 1 and 9) were asked from the children, independent of their actual performance.

First, it was found that there are LNS-knowers, i.e., several preschoolers know numbers up to 5, 6, 7 or 8 without knowing the next higher number in a GaN task. Our results are in line with former observations (Gunderson et al., 2015; Marchand & Barner, in press; Mussolin et al., 2012; J. B. Wagner & Johnson, 2011). These results are essential because one of the main initial arguments why children who know larger numbers than 4 are CP-knowers was that no LNS-knowers could be observed. Most probably, the reason why LNS-knowers were not observed in early studies is the limitations of the methods: Numbers larger than 6 were not measured and the samples were relatively small. Second, we found that number knowledge is not all-or-nothing but it is a gradual knowledge which property is not captured appropriately by the current GaN task versions. These results are also in line with former findings (Barner & Bachrach, 2010; Gunderson et al., 2015; O'Rear et al., 2020; Sarnecka & Gelman, 2004; K. Wagner et al., 2019). Both of these main results could be replicated: The same pattern was observed in two main studies and also in a preparatory study.

4.1 A refined version of the GaN task

These findings have several implications how the methodological details of the GaN task could be refined. First, while the current method stops measuring number knowledge beyond 5, it is possible that the observed operational LNS-knowers are also conceptual subset-knowers. Therefore, it is worth to measure larger numbers as well otherwise operational LNS-knowers and potentially conceptual subset-knowers cannot be identified (see the next subsection why this is essential in many studies). Second, because the number knowledge is not all-or-nothing and because there could be gaps in children's number knowledge, instead of using the titration method, a wider range of numbers should be measured. In a refined version, either all numbers in the planned range should be measured, independent of the low performance for a single value, or an adaptive method should not only measure up to the first unknown number, but the next 2, 3, or 4 numbers should be measured as well.

Calculating number knowledge. Third, a more appropriate index of number knowledge could be applied. However, currently, it is not straightforward to provide recommendations for alternative indexes. Here we offer a few consideration, although future studies may clarify what the critical properties of this number knowledge are and which index(es) can capture those properties. (a) One option is to use the largest known number as an index for number knowledge. However, this calculation method cannot consider the gradual change in number knowledge. (b) Another option is that a value for a specific performance, e.g. 50% could be estimated in a similar way as threshold is estimated with sigmoid curve fit. Nonetheless, the calculation of this value may be somewhat complicated. (c) Similar to the previous method, it is possible that performance for several values could be considered concurrently with an approximation utilizing Bayesian method, as suggested by

Negen & Sarnecka (2011). Importantly, this method requires a more precise model of number knowledge development which is not readily available in the literature yet. (d) As a fourth option, one could calculate the mean performance in the task, supposed that the same numbers are asked for all children, even if this would result in a longer procedure. Unfortunately, this method would make the comparison of studies difficult if different ranges of numbers are used in different studies. (e) As a fifth possibility, the number knowledge could be characterized by two numbers: a small number where the performance first leaves the reliable range of knowledge (e.g., the smallest unknown number), and the largest number before the performance becomes entirely random, e.g., reaches 0%). A critical problem with this method is that it could be more sensitive to noise than the previous two options, leading to a less reliable index. Beyond these options, there could be other possibilities to determine the number knowledge. At the moment it is hard to tell which calculation method would be the most appropriate to use, and future research can clarify the pros and cons of different possibilities.

To summarize, the present results suggest that larger numbers than 4 should also be measured when investigating the symbolic number knowledge of preschoolers, and it should not be supposed that number knowledge is all-or-nothing. Still, at the moment it is an open question how the number knowledge should be calculated.

4.2 Validating the conceptual knowledge of LNS-knowers

The present work made a first step to clarify the status of the LNS-knowers, and from an operational viewpoint it was demonstrated that the operational LNS-knower pattern is not impossible in children when measured with the GaN task. However, a next step could be the validation and to investigate if operational LNS-knowers are conceptual subset-knowers or conceptual CP-knowers. Such a first work demonstrated that both SNS-knowers and LNS-knowers can compare only the symbolic numbers that is in their number level, e.g., a 6-knower can compare correctly the 3 vs 5 number pair, but not the 7 vs 9 pair (Krajcsi & Reynvoet, in preparation). This suggests that similar to SNS-knowers LNS-knowers have a limited conceptual understanding of numbers, and they cannot use any values even in a simple symbolic comparison task. Additional studies with diverse methods are needed to investigate the conceptual status of operational LNS-knowers.

Relatedly, there is a series of works that contrast the performance of subset-knowers and CP-knower in various numerical tasks (e.g., Davidson et al., 2012; Le Corre & Carey, 2007; Lipton & Spelke, 2006; Sarnecka & Carey, 2008; Sella et al., 2017; Sella & Lucangeli, 2020). These works are considered as validation works in a sense that they investigate these children's conceptual understanding. A typical finding in these studies is that even CP-knowers lack the understanding of fundamental properties of numbers. However, these works categorize LNS-knowers as CP-knowers. If operational LNS-knowers are in fact conceptual subset-knowers, then these validation works may rely on an invalid categorization. Therefore, one reason why these works found repeatedly that some of the CP-knowers could not perform even the simplest numerical task is because these

assumed CP-knowers could actually be conceptual subset-knowers. It is the question of future works whether categorizing LNS-knowers differently may resolve this seeming contradiction and whether correctly categorized CP-knowers can solve simple numerical tasks competently, or whether there are additional factors why some CP-knowers cannot entirely understand the fundamental properties of numbers.

4.3 Theoretical implications

The present finding also has critical consequences for the current theories of symbolic number learning. While the development of symbolic number learning is described in detail, it is still unknown what representations underlie this development. Two frequently cited proposals suggest that one of the representations supposed to function already in infants can support numerical development (Carey, 2009; Gallistel & Gelman, 1992; Piazza, 2010; vanMarle et al., 2018). First, we discuss the Approximate Number System model and the implications of the present results on this model, then the Object Tracking System model together with the relevant implications.

4.3.1 Approximate Number System

The Approximate Number System (ANS) model suggests that numbers are handled by an evolutionarily ancient representation, which is similar to the representations handling physical features known in psychophysics, working according to Weber's law (Dehaene, 1992; Moyer & Landauer, 1967). The functioning of the ANS was demonstrated in infants (Feigenson et al., 2004; Izard et al., 2009), supporting the idea that infants can already perceive the numerical values of sets. The ANS is noisy and the noise also determines what values could be discriminated. The precision (or noise) of this representation develops with age, and smaller noise enables children to differentiate number pairs with smaller ratio: While newborn infants need at least three times the difference between two sets to differentiate them, adults can be sensitive to as small a difference as 15% (Halberda et al., 2008, 2008; Izard et al., 2009; Piazza, 2010). Relying on this model, one account of symbolic number learning proposes that subset-knowers understand their first values with the help of this approximate system, where the slow development of number knowledge is caused by the slow increase of the precision of the approximate system (Gallistel & Gelman, 1992; Piazza, 2010): A new number is only learned when the precision of the approximate system becomes appropriate to differentiate that new number from its larger neighbor. Later, when the children understand the cardinality principle, another unspecified representation starts to support symbolic number understanding.

The ANS model is not only capable of explaining the present results, but seemingly these results fit the ANS model even better than the previously assumed pattern. First, the ANS model could be in line with the presence of conceptual LNS-knowers, because in the model there is no inherent property that would explain why the transition between subset-knowers and cardinality principle knowledge should be around the number 4. So this more flexible transition could be more in line with the ANS account than the previously assumed 4-based limitation. Second, the gradual

performance to give specific numbers is in fact more appropriate prediction of the model than the previously assumed sudden change in the performance if it is supposed that partially known numbers are also supported by the ANS. (Note, however, that this property may not be testable in the GaN task: K. Wagner et al., 2019.)

However, while it is possible that the ANS model can generate the observed pattern, the current results are not conclusive, because other alternative mechanisms can also generate that pattern (see an example in the next subsection). Additionally, in the recent years, several arguments suggested that the ANS model might be inappropriate for explaining various numerical phenomena which may question its role in the development of subset-knowers as well. (a) Regarding preschoolers, although it was proposed that the development of Weber fraction can lead the learning of the first symbolic numbers (Piazza, 2010), it was also proposed that the ANS change cannot be observed before a new number or the cardinality principle is learned (Shusterman et al., 2016), or the extent of Weber fraction change in subset-knowers is too small to account for the change in symbolic knowledge (Krajcsi et al., submitted), and this range should be even larger if LNS-knowers should also be accounted for. (b) From a methodological viewpoint, it is debated whether ANS and exact number abilities are related (Leibovich & Ansari, 2016; Mussolin et al., 2016). There is an active ongoing debate in the literature about whether ANS sensitivity has been measured appropriately in the studies that found a relation between the ANS and other symbolic mathematical capabilities. For example, visual discrimination abilities or inhibition processes influence the to-be-measured ANS sensitivity (Gebuis & Reynvoet, 2012), or different methods to measure the ANS acuity in fact measure different constructs (Smets et al., 2014). Although there could be advanced methods available to measure ANS precision (DeWind et al., 2015), they are not used widely yet, and some issues, such as the influence of inhibition, still remains to be resolved. (c) More generally, it has also been proposed that all symbolic numerical effects that have been attributed to the ANS can be explained with a simple mechanisms consisting of a discrete symbolic network (Krajcsi et al., 2022).

Taken together, although the present results may support the role of the ANS in symbolic number learning, the current data may also be explained by other models, and other works debate whether the ANS can play an essential role in the initial learning of symbolic numbers.

4.3.2 Object Tracking System

The Object Tracking System (OTS) model is a visual model that proposes that this mechanism can track a few objects in the visual field (Trick & Pylyshyn, 1994). Importantly, the capacity of this system is limited: Healthy adults can track at most 4 objects simultaneously. In the numerical cognition literature, it has been demonstrated that preverbal infants can track the number of objects up to 3 or 4, and implicitly they can track even the simple arithmetical operations involved in the changes in the sets (Wynn, 1992a). These tasks, where the limit of the solution is the set size, and not the ratio of the values, should be solved with a system other than the ANS, and this could be the OTS that can track individual objects up to 3 or 4 (Feigenson et al., 2004; Trick & Pylyshyn, 1994).

OTS was not only proposed to be a source of the numerical abilities of infants, but also as an account for initial symbolic number knowledge. It was proposed that either the OTS may support directly the subset-knowers number knowledge or based on the OTS, children may form different size of set templates in their long-term memory to handle specific values, and these object file-based templates may be the root of the symbolic number knowledge of subset-knowers (Carey, 2004, 2009; Reynvoet & Sasanguie, 2016). Beyond number 4, this system is not able to handle the values, because the OTS can only handle 4 objects, and another unspecified representation will process larger numbers, when the cardinality principle is understood.

Existence of LNS-knowers may question the role of object file explanation in subset-knowers. If OTS directly supports the number knowledge of subset-knowers, then the OTS explanation cannot account for the knowledge of LNS-knowers, where 5, 6 or other larger numbers cannot be represented by the OTS that can track only objects up to 4 (Feigenson et al., 2004; Trick & Pylyshyn, 1994). However, one may suggest that it is not directly the OTS that supports the knowledge of subset-knowers, but the long-term-memory templates that are build up based on the OTS (Carey, 2009). Still, this modified version should explain how this long-term-memory template mechanism becomes independent from the OTS system after some point. Overall, it is an open question how the OTS model can account for the existence of the LNS-knowers.

5 Conclusion

To summarize, we found that (1) there are operational LNS-knowers who can not give all numbers that are in their counting list and whose response pattern is not different qualitatively from the response pattern of SNS-knowers, and (2) the preschoolers' performance is not all-or-nothing, but their number knowledge decreases gradually as the numbers increase. These findings suggest that the way children's number knowledge is measured should be reconsidered: In the Give a number task, instead of the titration procedure, a whole range of numbers should be tested, and measurement should not stop at number 5, but larger values should be used as well. However, it is not clear yet how the number knowledge of a child should be calculated. From a theoretical viewpoint, the present result may question the OTS explanation in its current form, and while the data may be in line with an ANS explanation, the same patterns can also be explained by alternative explanations, which latter conclusion is more in line with other phenomena and arguments that question the role of the ANS in symbolic number processing.

Further research is needed to refine the procedure and evaluation method for the GaN tasks and, essentially, additional validations are needed to find out whether LNS-knowers are conceptual subset-knowers or conceptual CP-knowers. More precise and extensive description of the development is essential to clarify the role of the competing models in the initial symbolic number learning.

6 Acknowledgments

We thank Lilla Hodossy, Krisztián Kasos and Petia Kojouharova for their comments on earlier versions of the manuscript.

7 References

- Almoammer, A., Sullivan, J., Donlan, C., Marušič, F., Žaucer, R., O'Donnell, T., & Barner, D. (2013). Grammatical morphology as a source of early number word meanings. *Proceedings of the National Academy of Sciences*, 110(46), 18448–18453.
<https://doi.org/10.1073/pnas.1313652110>
- Barner, D., & Bachrach, A. (2010). Inference and exact numerical representation in early language development. *Cognitive Psychology*, 60(1), 40–62.
<https://doi.org/10.1016/j.cogpsych.2009.06.002>
- Barner, D., Libenson, A., Cheung, P., & Takasaki, M. (2009). Cross-linguistic relations between quantifiers and numerals in language acquisition: Evidence from Japanese. *Journal of Experimental Child Psychology*, 103(4), 421–440.
<https://doi.org/10.1016/j.jecp.2008.12.001>
- Carey, S. (2004). Bootstrapping and the origin of concepts. *Daedalus*, 59–68.
- Carey, S. (2009). *The origin of concepts* (1st ed.). Oxford University Press, USA.
- Cheung, P., Slusser, E., & Shusterman, A. A. (2016). 6-month longitudinal study on numerical estimation in preschoolers. *Proceedings of the Cognitive Science Society*.
<https://pdfs.semanticscholar.org/a8b9/e015eb6710be668e6c9fa126f3676e73c4ce.pdf>
- Condry, K. F., & Spelke, E. S. (2008). The development of language and abstract concepts: The case of natural number. *Journal of Experimental Psychology: General*, 137(1), 22–38.
<https://doi.org/10.1037/0096-3445.137.1.22>
- 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.
[https://doi.org/10.1016/0010-0277\(92\)90049-N](https://doi.org/10.1016/0010-0277(92)90049-N)

- DeWind, N. K., Adams, G. K., Platt, M. L., & Brannon, E. M. (2015). Modeling the approximate number system to quantify the contribution of visual stimulus features. *Cognition*, 142, 247–265. <https://doi.org/10.1016/j.cognition.2015.05.016>
- Feigenson, L., Dehaene, S., & Spelke, E. S. (2004). Core systems of number. *Trends in Cognitive Sciences*, 8, 307–314.
- Gallistel, C., & Gelman, R. (1992). Preverbal and verbal counting and computation. *Cognition*, 44, 43–74.
- Gallistel, C., & Gelman, R. (2000). Non-verbal numerical cognition: From reals to integers. *Trends in Cognitive Sciences*, 4(2), 59–65.
- Gebuis, T., & Reynvoet, B. (2012). The interplay between nonsymbolic number and its continuous visual properties. *Journal of Experimental Psychology: General*, 141(4), 642–648. <https://doi.org/10.1037/a0026218>
- 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., Mazocco, M. M. M., & Feigenson, L. (2008). Individual differences in non-verbal number acuity correlate with maths achievement. *Nature*, 455(7213), 665–668. <https://doi.org/10.1038/nature07246>
- Izard, V., Sann, C., Spelke, E. S., & Streri, A. (2009). Newborn infants perceive abstract numbers. *Proceedings of the National Academy of Sciences*, 106(25), 10382–10385. <https://doi.org/10.1073/pnas.0812142106>
- Krajcsi, A. (2021). Follow-up questions influence the measured number knowledge in the Give-a-number task. *Cognitive Development*, 57, 100968. <https://doi.org/10.1016/j.cogdev.2020.100968>
- Krajcsi, A., Fedele, M., & Reynvoet, B. (submitted). *Approximate Number System cannot account for the number acquisition by subset-knowers.*

- Krajcsi, A., Kojouharova, P., & Lengyel, G. (2022). Processing Symbolic Numbers: The Example of Distance and Size Effects. In J. Gervain, G. Csibra, & K. Kovács (Eds.), *A Life in Cognition: Studies in Cognitive Science in Honor of Csaba Pléh* (pp. 379–394). Springer International Publishing. https://doi.org/10.1007/978-3-030-66175-5_27
- Krajcsi, A., & Reynvoet, B. (in preparation). *Subset-knowers can compare only the numbers they know*.
- 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/Revue Canadienne de Psychologie Expérimentale*, 70(1), 12–23. <https://doi.org/10.1037/cep0000070>
- Lipton, J. S., & Spelke, E. S. (2006). Preschool children master the logic of number word meanings. *Cognition*, 98(3), B57–B66. <https://doi.org/10.1016/j.cognition.2004.09.013>
- Marchand, E., & Barner, D. (in press). Assessing the knower-level framework: How reliable is the Give-a-Number task? *Cognition*. <https://cognitivesciencesociety.org/cogsci20/papers/0095/index.html>
- Mix, K. S., Sandhofer, C. M., Moore, J. A., & Russell, C. (2012). Acquisition of the cardinal word principle: The role of input. *Early Childhood Research Quarterly*, 27(2), 274–283. <https://doi.org/10.1016/j.ecresq.2011.10.003>
- 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>

- Mussolin, C., Nys, J., Leybaert, J., & Content, A. (2012). Relationships between approximate number system acuity and early symbolic number abilities. *Trends in Neuroscience and Education*, 1(1), 21–31. <https://doi.org/10.1016/j.tine.2012.09.003>
- Mussolin, C., Nys, J., Leybaert, J., & Content, A. (2016). How approximate and exact number skills are related to each other across development: A review☆. *Developmental Review*, 39, 1–15. <https://doi.org/10.1016/j.dr.2014.11.001>
- Negen, J., Sarnecka, B. W., & Lee, M. D. (2011). 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’Rear, C. D., McNeil, N. M., & Kirkland, P. K. (2020). Partial knowledge in the development of number word understanding. *Developmental Science*, 23(5), e12944. <https://doi.org/10.1111/desc.12944>
- 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>
- 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. <https://doi.org/10.3389/fpsyg.2016.01581>
- 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>
- Sarnecka, B. W., & Gelman, S. A. (2004). Six does not just mean a lot: Preschoolers see number words as specific. *Cognition*, 92(3), 329–352. <https://doi.org/10.1016/j.cognition.2003.10.001>
- Sarnecka, B. W., & Lee, M. D. (2009). Levels of number knowledge during early childhood. *Journal of Experimental Child Psychology*, 103(3), 325–337. <https://doi.org/10.1016/j.jecp.2009.02.007>

- 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. <https://doi.org/10.1016/j.cognition.2016.10.010>
- Sella, F., & Lucangeli, D. (2020). The knowledge of the preceding number reveals a mature understanding of the number sequence. *Cognition*, 194, 104104. <https://doi.org/10.1016/j.cognition.2019.104104>
- Sella, F., Slusser, E., Odic, D., & Krajcsi, A. (2021). The emergence of children’s natural number concepts: Current theoretical challenges. *Child Development Perspectives*, n/a(n/a). <https://doi.org/10.1111/cdep.12428>
- 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>
- Smets, K., Gebuis, T., Defever, E., & Reynvoet, B. (2014). Concurrent validity of approximate number sense tasks in adults and children. *Acta Psychologica*, 150, 120–128. <https://doi.org/10.1016/j.actpsy.2014.05.001>
- Trick, L. M., & Pylyshyn, Z. W. (1994). Why are small and large number enumerated differently? A limited-capacity preattentive stage in vision. *Psychological Review*, 101, 80–102.
- vanMarle, K., Chu, F. W., Mou, Y., Seok, J. H., Rouder, J., & Geary, D. C. (2018). Attaching meaning to the number words: Contributions of the object tracking and approximate number systems. *Developmental Science*, 21(1), e12495. <https://doi.org/10.1111/desc.12495>
- Wagner, J. B., & Johnson, S. C. (2011). An association between understanding cardinality and analog magnitude representations in preschoolers. *Cognition*, 119(1), 10–22. <https://doi.org/10.1016/j.cognition.2010.11.014>
- Wagner, K., Chu, J., & Barner, D. (2019). Do children’s number words begin noisy? *Developmental Science*, 22(1), e12752. <https://doi.org/10.1111/desc.12752>
- Wynn, K. (1990). Children’s understanding of counting. *Cognition*, 36(2), 155–193. [https://doi.org/10.1016/0010-0277\(90\)90003-3](https://doi.org/10.1016/0010-0277(90)90003-3)

Wynn, K. (1992a). Addition and subtraction by human infants. *Nature*, 358, 749–750.

Wynn, K. (1992b). Children's acquisition of the number words and the counting system. *Cognitive Psychology*, 24(2), 220–251. [https://doi.org/10.1016/0010-0285\(92\)90008-P](https://doi.org/10.1016/0010-0285(92)90008-P)

8 Supplementary materials

(This section will be a separate file, and it is included in the present file for the purpose of review.)

8.1 Methods

Three- and four-years-old preschoolers solved the Give a number task where numbers from 1 to 6 were asked three times. This study was a preparatory study for the main study presented in the manuscript. These results are presented to demonstrate the robustness of the findings.

The preliminary study used the same methods as the main study with the following exceptions. (1) In the Give a number task at the end of the trials no question was asked. (2) The numbers were asked between 1 and 6 (instead of between 1 and 9) in the following order: 1, 3, 2, 6, 4, 5. (3) The series of 6 numbers were asked 3 times, instead of 4 times.

In the analysis, the methods of the main study was applied. A number was still considered to be known if more than 50% of the time it was responded correctly, however, since numbers were asked only 3 times, 66% and 100% correct response rate was considered as the number is known, and 33% and 0% correct responses denoted that the number is not known reliably.

A session lasted approximately 5-7 minutes. Data were collected by three experimenters, all of them blind to the question of the study.

The study was approved by the ethics committee of the Cognitive Psychology Department, Eötvös Loránd University.

8.1.1 Participants

Seventy-four preschoolers participated in the study. Two children who could not complete the tasks were excluded from the analysis. The final sample consisted of 72 preschoolers, 40 girls and 32 boys, with a mean age of 46.5 months, SD of 5.3 months, ranging between 35 and 60 months. Seven Hungarian preschools were involved in the measurement, 6 of them in the capital, and 1 of them in a country town, all of them mostly receiving children of middle-class families.

8.2 Results and discussion

The raw data can be found at <https://osf.io/z28hd/>.

In the counting list knowledge task, 56 (77.8%) children could recite the counting list at least up to 6. In four other cases, the children could not solve the task, showing that their counting list knowledge was below 3 (these children were either pre-, 1- or 3-knowers). Two further preschoolers showed smaller counting list knowledge than their number knowledge (they were 4- and 6-knowers measured with the largest known number calculation method), potentially reflecting some reliability problem of the counting list task. The remaining 10 children knew the counting list up to 4 or 5.

Children not able to recite the counting list at least up to 6 may be limited in their performance in the Give a number task, however, because this limited counting list knowledge cannot influence the results regarding the potential gaps they were not excluded from that analysis. On the other hand, counting list knowledge is relevant while looking for 5-knowers, because a potentially CP-knower may not give 6 when the counting list is only known up to 5, therefore, counting list knowledge is reported in that analysis.

8.2.1 LNS-knowers

We found preschoolers who knew numbers up to 5, but did not know the number 6 (Figure 6). According to the titration-like calculation method, there were five 5-knower preschoolers in our sample (all of their counting list knowledge were higher than 6), and according to the largest known number calculation method, there were ten 5-knowers (only two of them had a limited counting list knowledge up to 6, therefore, most of them were not limited by their counting list knowledge). Inspecting the individual data (e.g., see participants with mean performance of 89% in Figure 5), this result does not seem to be a statistical artifact, but a reliable repeating pattern reflecting an actual limit in their knowledge.

8.2.2 Partial knowledge of numbers

The pattern of the detailed performance for all numbers in all preschoolers also replicates the finding of the main study: Performance decreases gradually in ranges of values instead of a sudden performance drop between two neighboring values (Figure 5).

Figure 5. Individual performance in the Give a number task. See Figure 1 caption for more details about the columns. Only a single child is shown out of the 9 perfectly performing children.

Overall, the preliminary measurement replicated and extended the results of the main study: (1) there are LNS-knowers, and (2) children's number knowledge is not all-or-nothing, but they have partial knowledge of numbers and gaps in their known number range.

As in the main study, the two number knowledge calculation method categorized differently some of the children (Figure 6). Twenty-five percent of the children (18 out of 72) were categorized differently by the two methods: Some of the 0-, 1-, 2- and 3-knowers measured with the classic method were categorized as 2-, 3-, 4-, 5- or 6-knowers with the new method (points above the $y = x$ line). (For the sake of simplicity, here we term children who successfully could give six as 6-knowers, although based on our tasks it is not possible to tell whether they are CP-knowers or subset-knowers.). It is not possible to tell whether there are gaps in the knowledge of the 5- and 6-knowers (as measured with the classic calculation methods), because 7 and 8 should have been asked to see if larger numbers are known after not knowing 6 and 7. Relatedly, it is possible that there could be gaps in the knowledge of the 4-knowers (measured with the classic calculation method), but in our sample there were only one 4-knower, which sample size may obviously be too small to find children with gaps in their knowledge reliably.

Figure 6. Number of children with a specific number knowledge calculated with the titration-like method (smallest unknown number – 1) vs. calculated with the largest known number method. Largest size denotes 18 children, smallest size denotes 1 children. (See also Figure 1. caption for more details.)