Cognition 223 (2022) 105019

Contents lists available at ScienceDirect

Cognition

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

Extending ideas of numerical order beyond the count-list from kindergarten to first grade

Jane E. Hutchison [a], Daniel Ansari [b], Samuel Zheng [c], Stefanie De Jesus [c], Ian M. Lyons [a,*]

[a] Georgetown University, Department of Psychology, United States of America
[b] University of Western Ontario, Department of Psychology and Brain & Mind Institute, Canada
[c] Toronto District School Board, Research and Development, Canada

ARTICLE INFO

Keywords:
Ordering
Count-list
Development
Kindergarten
First grade

ABSTRACT

Ordinal processing plays a fundamental role in both the representation and manipulation of symbolic numbers. As such, it is important to understand how children come to develop a sense of ordinality in the first place. The current study examines the role of the count-list in the development of ordinal knowledge through the investigation of two research questions: (1) Do K-1 children struggle to extend the notion of numerical order beyond the count-list, and if so (2) does this extension develop incrementally or manifest as a qualitative re-organization of how children recognize the ordinality of numerical sequences. Overall, we observed that although young children reliably identified adjacent ordered sequences (i.e., those that match the count-list; '2-3-4') as being in the correct ascending order, they performed significantly below chance on non-adjacent ordered trials (i.e., those that do not match the count-list but are in the correct order; '2-4-6') from the beginning of kindergarten to the end of first grade. Further, both qualitative and quantitative analyses supported the conclusion that the ability to extend notions of ordinality beyond the count-list emerged as a conceptual shift in ordinal understanding rather than through incremental improvements. These findings are the first to suggest that the ability to extend notions of ordinality beyond the count-list to include non-adjacent numbers is non-trivial and reflects a significant developmental hurdle that most children must overcome in order to develop a mature sense of ordinality.

1. Introduction

Symbolic numbers (e.g., Arabic numerals) hold both cardinal and ordinal meaning. Cardinality refers to the underlying quantities that number symbols represent (how many?), while ordinality refers to the ordinal relations amongst number in a sequence (what position?). While efforts to understand how humans represent and process symbolic numbers have long focused on the role of cardinality, the role of ordinality has only recently made its way to the forefront of research in numerical cognition (for a review, see Lyons, Vogel, & Ansari, 2016). Overall, this growing body of literature characterizes ordinal processing as a fundamental numerical skill that plays an important role in both the representation and manipulation of symbolic numbers in both children and adults (e.g., Lyons & Beilock, 2011; Lyons, Price, Vaessen, Blomert, & Ansari, 2014; Sasanguie, Lyons, De Smedt, & Reynvoet, 2017; Sasanguie & Vos, 2018). Despite the upsurge in research and recognition of its relative importance, major gaps persist in terms of our understanding of how ordinal understanding of number develops. What can perhaps be said most clearly is that, while children often have early exposure to ordered lists of numbers by learning to recite parts of the count-list ("one, two, three…"), children often do not understand the meanings of these numbers until much later (Wynn, 1992). Further, their understanding of ordinality is thought to lag behind that of cardinality in general (Colomé & Noël, 2012; Knudsen, Fischer, Henning, & Aschersleben, 2015; Michie, 1984; Spaepen, Gunderson, Gibson, Goldin-Meadow, & Levine, 2018). However, while the count-list appears to play a central role in early understanding of numerical order (Lyons et al., 2016) the precise nature of this influence is not known. Here, we examine how the count-list shapes – for better or worse – changes in how kindergarten and first-grade children process different kinds of numerical order.

1.1. Ordinality as a fundamental numerical skill

Understanding how numerical representations come to be over the course of children's development can inform millennia-old questions

* Corresponding author at: 306N White-Gravenor Hall, 37th and O Streets, NW., Washington, DC 20007, United States of America.
E-mail address: ian.lyons@georgetown.edu (I.M. Lyons).

https://doi.org/10.1016/j.cognition.2022.105019
Received 8 January 2021; Received in revised form 22 December 2021; Accepted 6 January 2022
Available online 1 February 2022

about the nature of concepts like 'number', 'quantity' and 'order'. A written number can represent cardinal value, ordinal relations between objects, or even ordinal relations between other numbers. Indeed, foundational work on the construction of coherent logico-mathematical systems do not always agree whether the central meaning of a number is its cardinal value (e.g., Bertrand Russel) or its ordinal position (e.g., Giuseppe Peano). Hence, investigating the constraints, limitations and advances children make in constructing basic numerical ideas – such as numerical order – has implications for our theoretical understanding of the basic cognitive foundations upon which more sophisticated mathematics are built (Lyons et al., 2016). Evidence supporting ordinality in particular as a fundamental numerical skill stems from a growing body of work linking ordinal processing to basic symbolic quantity processing, as well as to complex arithmetic performance.

First, a growing body of literature has documented robust and consistent associations between ordinal processing and more complex math skills such as arithmetic (Attout & Majerus, 2015; Goffin & Ansari, 2016; Lyons et al., 2014; Lyons & Ansari, 2015; Lyons & Beilock, 2011; Morsanyi, O'Mahony, & McCormack, 2017; Sasanguie et al., 2017; Sasanguie & Vos, 2018; Vogel et al., 2017). Specifically, the role of ordinality in arithmetic is thought to increase over developmental time (Lyons et al., 2014; Sasanguie & Vos, 2018). For example, in a cross-sectional sample of 1463 elementary school children, Lyons et al. (2014) observed that while magnitude comparison was the strongest predictor of arithmetic in first and second grade, the predictive contributions of ordinality increased across grade-level ultimately overtaking cardinality as the strongest predictor of arithmetic by the end of sixth grade. In line with this, Sasanguie and Vos (2018) observed that while cardinal processing mediated the relation between number ordering and arithmetic in first grade, ordinal processing mediated the relation between magnitude comparison and arithmetic in second grade. Together, these findings point toward a developmental shift in the early years of formal schooling in which knowledge of symbol-symbol associations replaces knowledge of symbol-quantity associations as a crucial mechanism in the development of sophisticated math skills.

Going beyond correlations, Merkley (2015) trained two groups of 6-year old children to use a set of abstract symbols in a numerical context. The first group was provided with only cardinal information when learning the symbols, whereas the second group was provided with both cardinal and ordinal information. Merkley observed that children learned to use the newly-acquired symbols in a numerical context (determine which of two is larger) only if they were provided with *both* ordinal and cardinal information during the training. Those who were provided with only cardinal information during the training were at chance on the comparison task. Similar findings have been observed in adults (Lyons & Beilock, 2009; Merkley, Shimi, & Scerif, 2016). Overall, these findings suggest that understanding the ordinal relations between numbers may play a crucial role in tapping the full potential of what a system of abstract symbolic numbers can do.

More broadly, it is useful to point out that number symbols are not natural objects. As such, they can help us think beyond the constraints of our immediate perceptions, but they therefore pose a major set of conceptual challenges. For instance, there is evidence suggesting that acquiring the ability to understand number symbols may alter how one processes (nonsymbolic) perceptual magnitudes (e.g., Lyons, Bugden, Zheng, De Jesus, & Ansari, 2018; Piazza, Pica, Izard, Spelke, & Dehaene, 2013), and this effect may be strongest for larger magnitudes where the parallels between concrete objects and abstract quantities begin to break down (Hutchison, Ansari, Zheng, Jesus, & Lyons, 2020). Successfully learning to disentangle number symbols from concrete perceptions can allow us to understand not only that 1,000,001 is greater than 1,000,000 without ever having directly perceived either magnitude, but also that the difference between those two quantities is exactly the same as that between far more quotidian quantities like 1 and 2 (Núñez, 2017). Understanding numerical order is key to the insight noted above. Indeed, a fundamental aspect of ordinality is that it allows one to move *beyond* the count-list to form a rich network of associations amongst non-adjacent series of numbers as well. Recent work, however, suggests that extending one's notion of order beyond the count-list may not be trivial – a topic we turn to in the next section.

1.2. *The role of the count-list in the development of ordinality*

Young children gain practice ordering symbolic numbers (in the form of spoken words) well before developing an understanding of what those symbols mean. Specifically, young children are first introduced to an ordered sequence of numbers when learning to count. Most young children are able to recite the count-list in the correct order by the age of 2 or 3, before they have acquired the meaning of each number word (Wynn, 1992). While the ability to recite the count-list early on does not reflect a deep understanding of the ordinal associations amongst numbers, it suggests that from a young age, children are committing the count-list to memory. It thus seems reasonable to hypothesize that the count-list plays a key role in shaping later understanding of numerical order (Lyons et al., 2016).

Recent work by Gilmore and Batchelor (2021) investigated whether verbal counting skills could explain the robust relation between numerical ordering tasks and more complex math tasks such as mental arithmetic. Contrary to what Lyons and Ansari (2015) observed, Gilmore and Batchelor found that it did. An important difference between the two studies is that Lyons and Ansari measured very simple rote counting up to 9, whereas Gilmore and Batchelor measured more complex counting, such as 'counting-on' and 'counting-down' from a higher number (e.g., 25). Together, these results suggest that complex sequential understanding may be especially important in the development of the numerical foundations of children's mathematical understanding. It remains unclear, however, just how straightforward extending notions of numerical order beyond the standard integer count-list may be.

In particular, it is unclear whether children's extensive experience with the count-list facilitates or perhaps even hinders their ability to expand their notion of numerical order beyond this oft-recited list. Both adults and children are slower and less accurate when identifying non-adjacent, ordered sequences of numbers (e.g., 2–4-6) as 'in-order' compared to sequences like 3–4-5 (Franklin & Jonides, 2009; Gilmore & Batchelor, 2021; Lyons & Ansari, 2015; Lyons & Beilock, 2013; Turconi, Campbell, & Seron, 2006). One simple explanation for this phenomenon (commonly referred to as the reverse distance effect, RDE) is that repeated practice with the count-list facilitates processing of count-list sequences. However, Gattas, Bugden, and Lyons (2021) directly tested this assumption and found it lacking. It is less that processing sequences like 3–4-5 is facilitated, and more that processing sequences like 2–4-6 is *impeded*. When participants were forced to classify sequences like 3–4-5 as '*not*-in-order', performance was largely unaffected. However, when participants were forced to classify sequences like 2–4-6 as '*not*-in-order' their performance substantially *improved*. If participants have a strong bias to 'see' count-list sequences as in-order, then forcing them to indicate the opposite should be highly challenging; but it was not. Conversely, if participants have a strong bias to 'see' non-count-list sequences as 'not-in-order' (even when they actually are), then permitting them to indicate as much should be relatively easy – which was exactly what Gattas and colleagues found. Further, the authors found that this latter phenomenon largely explained the RDE in ordinal verification tasks. Note that this work was conducted with highly numerate adults (university undergraduates), so it would appear that the legacy of the count-list in numerical order processing may be less to facilitate notions of what *is* in order and more to create a lasting set of biases about what is *not* in order. Gattas et al. speculated that their observations were vestiges of an earlier developmental struggle to extend notions of numerical order beyond the count-list, but they were of course limited to speculation due to their adult sample.

Interestingly, in line with the results discussed above, Gilmore and

Batchelor (2021) reported a bimodal distribution in the extent to which children in their study (ages 6.0–8.9 yrs) appeared to understand that non-count-list sequences (e.g., 2–4-6) should be classified as 'in-order'. Roughly a third (41 of 62) of the children in their sample performed at or below chance (50%) when classifying distance-2 (e.g., 2–4-6) or distance-3 (e.g., 1–4-7) trials as 'in-order'. Indeed, about a quarter of the sample (15 children) responded exclusively (0% accuracy) that such trials were 'not-in-order'. This set of results potentially lends some credence to the speculation noted above by Gattas et al. (2021). However, it is important to note several limitations. First, Gilmore and Batchelor's sample was relatively small ($N = 62$), comprised a relatively wide range of ages (the youngest of which was 6 years) and it was cross-sectional. This is not meant as a criticism of the Gilmore and Batchelor study as a whole, as these shortcomings do not speak to the central conclusions of their paper. Rather, if, in the current context, one wished to probe more deeply *when* and *how* children learn to extend their notions of numerical order beyond the integer count-list, then one might do well to investigate a larger sample of children starting at a younger age and following them across multiple time-points. In particular, it would be useful to understand whether it is the majority of children, or only a small subset, who struggle to extend the notion of numerical order beyond the count-list at the beginning of formal schooling (start of kindergarten). Further, it would be helpful to know roughly when this extended understanding of numerical order begins to emerge, and whether the transition is gradual or a reflection of a qualitative shift in children's thinking. Finally, knowing the answer to the above may permit a preliminary investigation into the factors that may help children learn to go beyond the count-list.

1.3. Current study

Overall, ordinal processing is fundamental to both the representation and manipulation of symbolic quantities. As such, it is important to understand how children come to develop a sense of ordinality in the first place. The current study longitudinally examined whether the count-list influences how children at the outset of formal education (kindergarten and 1st-grade; ages 5–6 years) recognize different kinds of numerical order (sequences that do and do not match the count-list). We chose this age range for several reasons. First, by kindergarten, the majority of children can recognize written number symbols, recite the count-list, and demonstrate a complete understanding of cardinality (e.g., Colomé & Noël, 2012; Jordan, Kaplan, Oláh, & Locuniak, 2006; Wynn, 1992). As such, we could reasonably assess ordinal recognition using a version of the standard ordinality verification task popular elsewhere in the literature. This in turn allowed us to test for the existence of the RDE in children younger than 1st-grade. Finally, as they are encountering formal math schooling perhaps for the first time, we anticipated this age range might be apt to show changes in children's understanding of numerical order, especially with respect to extending one's sense of order beyond the count-list. Hence, we investigated (1) whether K-1 children struggle to extend the notion of numerical order beyond the count-list, and (2) whether this extension develops incrementally or manifests as a qualitative reorganization of how children recognize the ordinality of numerical sequences. Further, identifying when the above development begins to occur would allow for a post hoc assessment of which basic numerical capacities predict (and thus potentially contribute to) the expansion of children's ordinal understanding.

To address whether K-1 children struggle to extend the notion of numerical order beyond the count-list, we used a standard ordinality recognition task and examined performance on trials that either did match the count-list (adjacent numbers, '3–4-5') or did not (non-adjacent numbers; '2–4-6'). Given prior research documenting an RDE during ordinal processing in both adults (Franklin & Jonides, 2009; Lyons & Ansari, 2015; Lyons & Beilock, 2013; Turconi et al., 2006) and children (Lyons & Ansari, 2015), we hypothesized that children would perform worse on non-adjacent compared to adjacent trials, starting as early as the beginning of kindergarten. If so, an important follow-up question is to begin to understand the nature and extent of underperformance on non-adjacent trials like '2–4-6'. Are kindergarteners only a little less proficient on these trials relative to adjacent trials (i.e., still well above chance performance), or do some, or perhaps even the majority of children, demonstrate a more fundamental misunderstanding of ordinality, as evidenced by performance at or below chance on these trials? The latter situation would suggest that a given child at this age indeed struggles to understand that the notion of numbers being 'in-order' applies to sequences beyond the integer count-list.

The second question we investigated was how children's understanding of different types of numerical order potentially develops across three time-points: fall and spring of the kindergarten school-year into 1st-grade. In general, we expected to see overall performance on the order verification task improve with age, as has been previously shown in other studies (Lyons & Ansari, 2015). Crucially, however, we examined whether the different trial-types – especially adjacent (e.g., '3–4-5') and non-adjacent (e.g., '2–4-6') trials show different developmental trajectories. Of particular interest here is the development of non-adjacent trials. Do age-related improvements on these trials occur incrementally, or more suddenly? If the latter, this might provide further indication that extending one's sense of order beyond the count-list comprises a qualitative – or perhaps even conceptual – shift in children's understanding of the greater range of what constitutes numerical order. Proposing a conceptual-level change in understanding is a rather strong claim, and so the current study probes this question in two different ways.

First, we examined how the underlying distributions in performance changed over time. If developmental change in performance on non-adjacent ('2–4-6') trials reflects simple incremental change, we would expect to observe a unimodal distribution at all time-points with the peak (mode) gradually shifting in a positive direction (i.e., to the right in a standard histogram). Alternatively, individual children may experience a kind of 'eureka' moment that involves a qualitative re-classification of non-adjacent trials from 'not in-order' to 'in-order'. In this case, we should see a bimodal distribution, reflecting two sub-populations of children – those who 'get' that numerical order extends beyond the count-list, and those that don't. Crucially, this latter hypothesis applies only to non-adjacent trials, and not to the other trial-types because the 'eureka' moment described above should not alter how the other types of trials are classified in the ordinal verification task.

A second, perhaps more formal method of testing whether changes in performance on non-adjacent trials undergo a gradual or abrupt shift is to investigate changes in the intercorrelations amongst the different trial types over time. In total, the ordinal verification task comprises four trial-types: in-order adjacent (e.g., '3–4-5'), in-order non-adjacent (e.g., '2–4-6'), mixed (or not-in-order) adjacent (e.g., '3–5-4'), and mixed non-adjacent (e.g., '2–6-4'). One can examine the correlations between trial-types, and how these inter-correlations do or do not change over development. In the case of a major change in how children conceptualize one (and only one) of the trial-types amongst a large percentage of the sample, one would expect a dramatic change in how performance on that trial-type relates to performance on the other trial-types. In particular, if children abruptly switch from answering 'no, not in order' to 'yes, in order' for non-adjacent trials like '2–4-6' (but no such switch occurs for the other three trial types), then one should see a large change in the correlations between performance[1] on '2–4-6' trials and the other three trial-types. Importantly, this change in correlations should occur only for the subset of children who indeed made the above switch, and only across measurement time-points that straddle when that switch occurred. In conjunction with the more qualitative unimodal/bimodal

[1] Where performance is indicated by accuracy, and the criterion for an accurate response remains unchanged – as is the case in the current study.

test described above, this correlational approach can help unpack what may be happening within each of those sub-distributions of children.

Overall, to our knowledge, the current study is the first to examine in detail whether and how the count-list influences changes in how young children recognize different kinds of numerical order (namely, sequences that do and do not match the count-list). First, we investigated whether children in the early years of formal schooling struggle with extending notions of ordinality beyond the count-list (Question 1). Specifically, we tested whether children perform above chance on numerical order judgments that are in fact numerically in order, but that do not match the count-list (e.g., 2–4-6). Second, we investigated whether the ability to do so develops incrementally or as the product of a more abrupt qualitative shift (Question 2). This we did in two ways, (2a) by examining changes in performance distributions, and (2b) by examining changes in the interrelations between different types of ordinality judgments.

2. Methods

It is important to note that the data presented here come from a large, longitudinal data set, a portion of which has been described and reported elsewhere (Hutchison et al., 2020; Lyons et al., 2018). However, the current study addresses a unique set of theoretical questions and analyzes the data in a manner distinct from those reported previously.

2.1. Participants

Data were collected at three time points (fall senior kindergarten,[2] spring senior kindergarten, spring first grade) from 424 children across 35 schools within the Toronto District School Board (TDSB). Of these 424 students, 40 were removed due to missing data in one or more conditions of interest for at least one of the three time points. The final analysis sample consisted of 384 children (169 female; 34 not born in Canada). Mean age at the first time point (fall of the senior kindergarten year) was 5.18 years (range: 4.67–5.77; SD: 0.29). Socio-economic status (SES) was not available at the child level, although, it could be estimated for each school.[3] Schools were categorized as 0 = Low-SES (23.44%), 1 = Medium-Low-SES (31.25%), 2 = Medium-High-SES (33.33%), and 3 = High-SES (11.98%).

2.2. Procedure

2.2.1. Research collaborations

The data reported here are part of a joint research project between the TDSB and the University of Western Ontario (UWO), which was approved by the TDSB's External Research Review (ERRC). All data collection was conducted in collaboration with teachers, early childhood educators (ECEs) and administrators in TDSB schools. The Board authorized TDSB's Research and Development Department to collect assessment data and personal information for the purposes of the Board's educational planning. Parents/guardians of participating students were informed that classroom educators would be collecting the assessment data and that confidential student-level data would be kept within the TDSB's Research and Development Department. The TDSB's Research and Development Department was authorized to share depersonalized data (stripped of any school or student identifiers) with related research partners for this study. Assessment materials were approved by the University of Western Ontario's Non-Medical Research Ethics Board.

2.2.2. Data collection

Data were collected by the teachers and ECEs of the classrooms in which the testing took place. Teachers and ECEs were trained on administering the Numeracy Screener during an in-service work day. Administration of the Numeracy Screener was conducted during 15–20 min one-on-one testing sessions with the teacher/ECE and the student in a separate, quiet area at three time points: fall of senior kindergarten (2014), spring of senior kindergarten (2015), spring of first grade (2016). The average interval between the first two assessments was 191.73 days ($SD = 14.25$). The average interval between the second and third assessment was 382.04 days ($SD = 10.16$).

In each testing session, the teacher/ECE went over a predefined set of instructions with the student. Task-specific instructions and general guidelines were printed in the booklet on the page before the start of each task (see Appendix A). Before each task, the teacher/ECE went through several example items with the child to ensure that they understood the task (see below for specific instructions). After going over the instructions and practice items, the teacher/ECE started the timer, and the child began the task. Note that teachers/ECEs completed intensive in-service training to ensure they were aware of how to administer the screener, as well as the importance of adhering to all instructions, procedures and guidelines to help ensure data integrity.

2.2.3. Numeracy screener

The Numeracy Screener booklets were based on a design originally developed by Nosworthy, Bugden, Archibald, Evans, and Ansari (2013). The booklets contain six basic numerical tasks, although only one (number ordering) is of relevance for the current study.

2.2.4. Number ordering task

Verbatim instructions were as follows. “In this task, your job is to decide whether the three numbers are in the correct order (from left to right). If the numbers are in order, draw a line through the [check mark]. If the numbers are not in order, draw a line through the [X mark]. Let's practice the example problems here. [The child then completed 4 practice items – see below for additional details.] There will be problems on the backs of the pages as well. Make sure you don't skip any problems. You should try to complete as many problems as you can. You have 2 minutes. Work as fast as you can without making too many mistakes. If you do make a mistake, draw an X through the mistake and put a new line through the right answer.”

Children completed 4 practice items. Two items were ‘correct’, meaning children should mark the ✓ mark (1–2–3, 5–7–9), and two were ‘incorrect’, meaning children should mark the ✘ mark (1–3–2, 6–4–8). For this task, experimenters were also instructed that it was especially important to work through all example problems, providing feedback as needed to ensure that students understood what it meant for numbers to be in the ‘correct’ order:

> “During practice items, point out and explain any mistakes. Start with the top-left item and work across then down (the student should do this on the test items as well).
> “For this task, it is especially important to work through the example problems to ensure that students understand what it means for numbers to ‘be in order’. *The left two examples are in order; the right two examples are not.*”

A copy of the complete script and guidelines experimenters followed can be found in Appendix A. Note that in giving feedback on the practice

[2] In many Canadian provinces, kindergarten is split into ‘Junior’ and ‘Senior’ Kindergarten. Junior Kindergarten is similar to what is sometimes referred to as ‘preschool’ elsewhere, as it is often relatively informal in overall structure and available to children who are 4 years old. Senior Kindergarten is more similar to what is referred to as kindergarten elsewhere. Senior Kindergarten tends to be more formally structured and involves the instruction of basic formal concepts in mathematics and other areas.

[3] School SES was estimated from median family income, percentage of families below the Low-Income Measure, percentage of families on social assistance, percentage of parents without a high school diploma, percentage of parents with at least one university degree, and percentage of single-parent families.

items, experimenters were permitted to refer to numerical concepts such as 'increasing', 'going up' or 'getting bigger'. Importantly, experimenters were instructed *not* to provide a detailed explanation of what constitutes numerical order, or to provide specific strategies such as skip-counting techniques. This is because our aim was to investigate whether children could recognize different forms of numerical order, not to teach them how to do so (i.e., this was an observational, not an intervention study). Each of the above points was emphasized during the in-service training attended by all experimenters, and was subsequently re-emphasized throughout the process by the cooperating school board.

The main number ordering task was made up of 48 items, with 12 items per page. Each item consisted of three Arabic numerals (1–9) presented side-by-side in a rectangular grid (see Fig. 1). The child was asked to indicate whether the three numbers were in the correct order by either crossing through the ✓ mark (correct response for in-order trials) or the ✗ mark (correct response for mixed-order trials), which were presented below the number series. Of the 48 items, 14 consisted of in-order adjacent sequences (numbers that were separated by a numerical distance of 1; abbreviated as OA; see Fig. 1a), 10 consisted of in-order non-adjacent sequences (numbers that were separated by a numerical distance of 2; ONA; Fig. 1b), 14 consisted of mixed-order adjacent sequences (abbreviated as MA; Fig. 1c) and 10 consisted of mixed-order non-adjacent sequences (abbreviated as MNA; Fig. 1d). The full trial list can be found in Appendix B. Participants completed as many items as possible within 2 min.

Importantly, trial order was balanced such that, regardless of how far children progressed, they completed a similar proportion of the four trial types (OA, ONA, MA and MNA). Trial order was also balanced in terms of average numerical size (see Appendix B).

2.2.5. Task scoring

Raw scores were calculated as the total number of correct responses within the two-minute time limit. However, because children inevitably completed different numbers of trials within the two-minute time-frame, it is important to correct for guessing strategies. For instance, the average number of trials increased with age: T1: 21.2 trials attempted, T2: 22.0 trials completed, T3: 27.0 trials attempted. Hence, even if all children were randomly guessing at all time-points, one would expect to see raw scores increase (from 10.6 to 11.0 to 13.5). In addition, the analyses addressing Question 1 hinge critically on comparing performance against what would be predicted by chance (or random guessing).

Therefore, scores were corrected for guessing (a child who randomly guessed on all 48 items would have received a score of 24) in a manner agnostic to the number of completed trials (provided this was greater than 0). We did so via the formula $A = C - [I/(P-1)]$, where A is the adjusted score, C is the number correct, I is the number incorrect, and P is the number of response options (Rowley & Traub, 1977). Using this adjustment, random guessing yields an average adjusted score of 0, regardless of the number of trials completed. For example, on a 4-item multiple-choice exam (where each choice is equally probable), those who randomly guessed on 20 items would, on average, receive a raw score (C in the equation above) of 5. Therefore, those who used a guessing strategy in this example would receive an adjusted score (A) of: $5 - [15/(4-1)] = 0$. In the current study, the task items only had two alternatives with equal probability of being correct so the equation for the adjusted score is essentially the number of correct responses minus the number of incorrect responses ($A = C - I$).

An additional note is that this approach implicitly controls for individual variability in speed-accuracy trade-offs. This is because the total amount of time-on-task (2 min) was fixed across participants, and so the number of trials completed is an implicit measure of average response-time. In sum, adjusted scores account for different numbers of trials completed and account for speed-accuracy trade-offs, making this approach desirable on a number of fronts. Hence, adjusted scores were used in all analyses. Adjusted scores were calculated separately for the four conditions (OA, ONA, MA, MNA).

2.3. Task descriptions for post-hoc analyses

Three tasks were used as additional predictors in post hoc analyses: numeral comparison (NC), dot comparison (DC) and mixed-format comparison (MC). All three tasks were collected at time-point 2 (T2) at the same time that the ordinal verification task was collected. Each of the three tasks was presented with a 2-min time-limit in the same manner as the ordinal verification task. Each comparison task comprised 72 total items, with 12 items per page. Prior to starting the timer, the experimenter provided instructions and worked through four example trials. Brief descriptions of each task are provided below. Note that these data overlap with data used in Lyons et al. (2018) and Hutchison et al. (2020), though the specific analyses conducted and hypotheses tested here are distinct from those publications. Scoring for these tasks, including use of adjusted scores, was computed in the same manner as the ordinal verification task (described in the previous section).

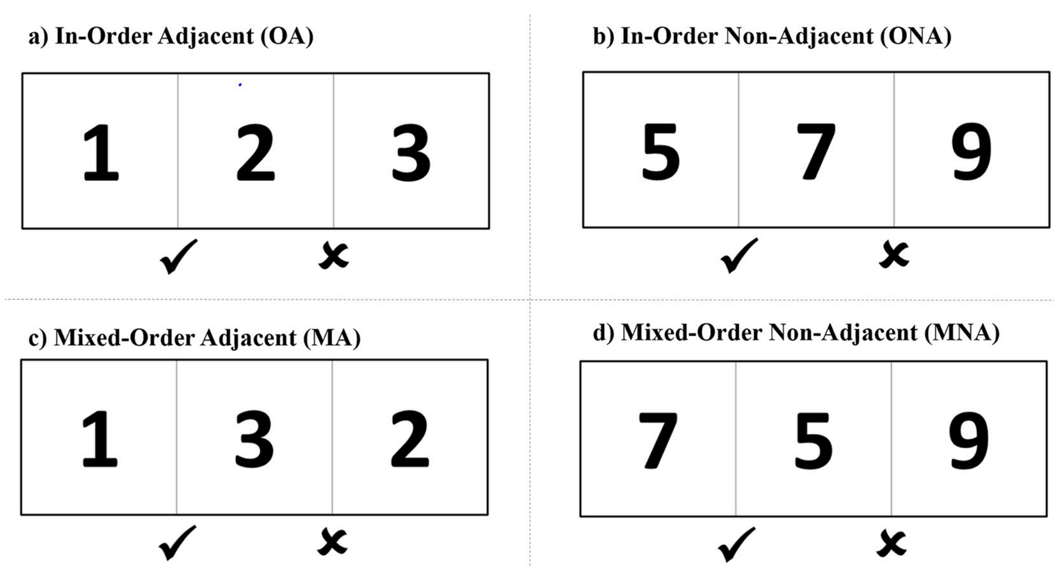


Fig. 1. shows examples of all trial types for the ordering task.

2.3.1. Numeral comparison (NC)

Children were told, "In this task, your job is to decide which of the two numbers is bigger. Draw a line through the box with the number that means the most things." Numerals ranged from 1 to 9 with absolute numerical distances $|n_1 - n_2|$ of 1 to 3, and ratios (min/max) from 0.250 to 0.889. Specifically, all 15 combinations of 1–9 with distances of 1 or 2 were included, along with 3 combinations with distance 3 ({1,4}, {3,6}, {6,9}). This yielded 18 possible combinations. Of these 18, 9 were permuted such that the larger number was on the left, and the other 9 were permuted such that the larger number was on the right. The 9 trials were chosen such that the larger side was in no way related to numerical size, distance or ratio. The next 18 trials were arranged in the opposite manner. The last 36 trials were determined in the same manner. Trial order was then pseudo-randomized within each set of 18 trials such that, for any n^{th} item in the sequence, average numerical ratio, size and distance were equated across comparison tasks (numeral, dot, mixed). This final step ensured that, if, for instance, a given child completed exactly 10 trials on each of the three comparisons tasks, the ratios (or sizes or distances) encountered on each task would not have differed significantly across tasks (all $ps > 0.20$). In other words, comparing performance across tasks was not confounded with these numerical factors.

2.3.2. Dot comparison (DC)

Children were told, "In this task, your job is to decide which of two boxes contains more dots. Draw a line through the box that has the most dots in it." Children were also instructed, "Don't try to count the dots. Instead, just look at the dots and try your best to guess which side has more dots in it." Numerosities and trial order were determined in the same manner as the numeral comparison task (described above). In addition, two versions of a given permutation were created. In one version, dot area was positively correlated with numerosity, and overall contour length was negatively correlated with numerosity; in the other version, the opposite was true. On a given trial, the two parameters were thus in opposition; between trials, relying on any single parameter would have led to chance performance (Gebuis & Reynvoet, 2012). Parameter version order was further pseudo-randomized such that it was not informative of the correct answer within a given segment of trials.

2.3.3. Mixed-format comparison (MC)

Children were told, "In this task, your job is to decide whether a number or a group of dots means more things. If the number means more things, draw a line through the number. If the dots mean more than the number, then draw a line through the dots." As with dot comparisons, children were also instructed, "Don't try to count the dots. Instead, just look at the dots and try your best to guess which side means more." Numerosities and trial order were determined in the same manner as the numeral comparison task (described above). In addition, which side contained the numeral and which the dots was pseudo-randomized such that it was not informative of the correct answer within a given segment of trials.

2.4. Data availability

Data are freely available for download at [Open Science Framework link to be provided should manuscript be accepted for publication].

3. Results

3.1. Question 1: do K-1 children struggle to extend notions of ordinality beyond the count-list?

Our first research question asked whether K-1 children struggle to extend the notion of ordinality beyond the count-list. To do so, we examined trial-based performance on a standard ordering task with ordered trials that did match the count-list (e.g., OA, '3–4-5') and ordered trials that did not (e.g., ONA, '2–4-6) as the primary trial-types of interest. Mean performance for each condition at each time-point is summarized in Fig. 2. Across all three time-points (T1, darkest bars in Fig. 2), children performed above chance (adjusted scores >0) on three out of the four conditions (OA, MA, MNA) [all $ps < 0.001$, all $ds > 0.66$]. Comparatively, children performed significantly *below* chance in the ONA condition across all three time-points (adjusted scores <0; blue bars in Fig. 2 [T1: $t(383) = -17.02$. $p < .001$, $d = -0.87$; T2: $t(383) = -14.07$. $p < .001$, $d = -0.72$; T3: $t(383) = -3.93$. $p < .001$, $d = -0.20$].

Thus, children on average reliably classified in-order trials that did not match the count-list as 'not in-order'. Note that this did not extend to adjacent trials, as children reliably identified OA trials (in-order trials that do match the count-list) as 'in-order' across all three time-points. Moreover, the fact that children performed above chance on average at all ages in three of the four conditions (OA, MA, MNA) indicates children understood and could complete the central demands of the task; children's difficulties were thus largely limited to the ONA trials.

3.2. Question 2: does the ability to extend notions of ordinality beyond the count-list develop incrementally or qualitatively?

Our second research question investigated how children's understanding of order develops across three time-points: fall of kindergarten, spring of kindergarten, and spring of first grade. Specifically, we were interested in whether the different trial-types, especially in-order adjacent (e.g., OA, '3–4-5') and non-adjacent (e.g., ONA, '2–4-6') trials, show different developmental trajectories.

Performance significantly improved from each time-point to the next for all three conditions [all $ps < 0.001$, all $ds > 0.20$]. In contrast, mean performance in the ONA condition did not change at all from the beginning to the end of kindergarten [from T1 to T2: $t(383) = 0.08$, $p = .936$, $d = 0.004$, remaining below chance at T2 [$t(383) = -14.07$, $p < .001$, $d = 0.72$]. Performance did, however, increase on ONA trials from T2 to T3 [$t(383) = 5.05$, $p < .001$, $d = 0.26$]. Notably, despite the significant increase, mean performance in the ONA condition remained significantly below chance by the end of 1st grade [T3: $t(383) = -3.93$, $p < .001$, $d = 0.20$]. Thus, while performance improved, children still tended to classify all in-order non-adjacent trials as 'not in order', even by the end of 1st grade.

While average performance on ONA trials remained below chance by the end of first grade, it did improve significantly from T2 to T3 (Fig. 2, medium to light blue bars). This improvement could either be indicative of incremental change in which the majority of children are beginning to slowly improve over time, or it could reflect a more abrupt change in which a subgroup of children shift from perceiving in-order non-adjacent sequences as 'not in-order' to 'in-order'. The latter would provide indication that extending one's sense of order beyond the count-list requires a conceptual shift in children's ordinal understanding. We attempted to differentiate between these two alternatives (incremental versus conceptual change) using both qualitative and quantitative methods. First, we examined change in the underlying performance distributions for each trial-type over time (qualitative). Second, we investigated change in the intercorrelations amongst the different trial types over time separately for those who perceive in-order non-adjacent (ONA) trials as 'in-order' by the end of first grade and those who do not (quantitative).

3.2.1. Research Question 2a – changes in the underlying performance distributions

If the improvement on ONA trials is indicative of incremental change, we would expect to observe a unimodal distribution of performance on this trial type at all three time points with the peak (mode) increasing (shifting to the right) from T2 to T3. Conversely, if the improvement is indicative of a qualitative re-classification of non-adjacent ordered trials from 'not in-order' to 'in-order', we should see a bimodal distribution, reflecting two sub-populations of children –

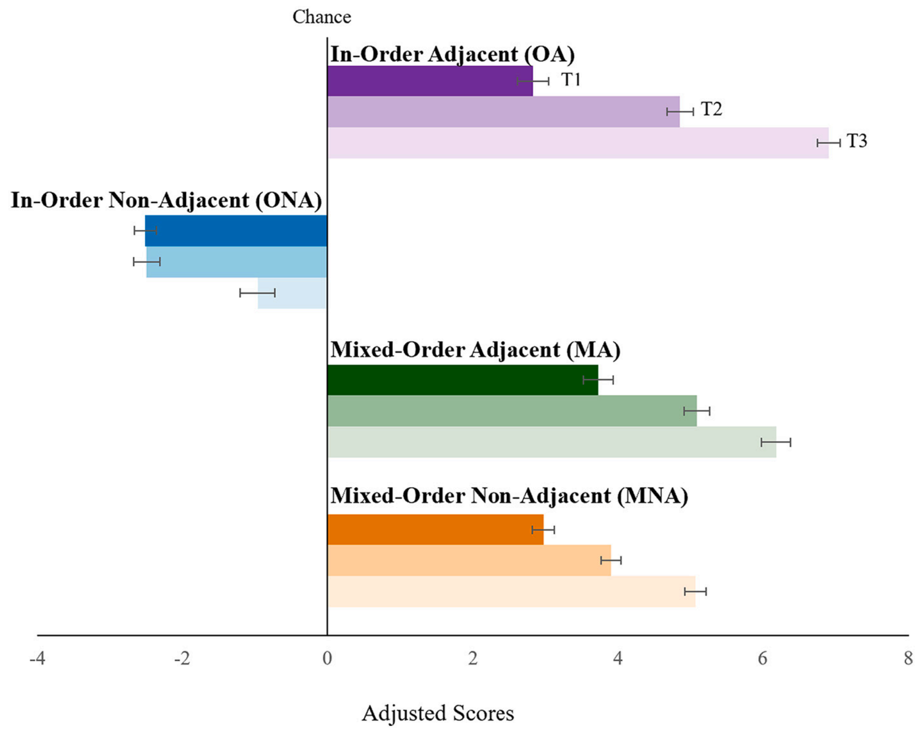


Fig. 2. shows mean performance on the ordering task, broken down by condition and time-point. Scores are adjusted so that a score of 0 indicates chance performance. Error bars reflect standard errors of the mean. T1 = beginning of kindergarten, T2 = end of kindergarten, T3 = end of 1st grade.

those who 'get' that numerical order extends beyond the count-list, and those who don't. Importantly, because we saw steady increases in the other three conditions (OA, MA and MNA) across all three time-points, we would expect to see a rightward-shifting unimodal distribution at all three time-points for these conditions.

As can be seen in Fig. 3, the distribution of performance on ONA trials at the beginning and end of kindergarten (T1 and T2, respectively) was unimodal with a peak centered around an adjusted score of approximately −5, indicating that most children in the sample fell below chance at this stage in development. However, by the end of first grade (T3), the distribution of performance on ONA trials became bimodal, with one peak centered over an adjusted score of approximately +5 and another centered over an adjusted score of approximately −5. This bimodality was specific to ONA trials, as performance on the other three trial types (OA, MA and MNA) followed a unimodal distribution across all three time points. These data indicate that a subgroup of children shifted from perceiving in-order non-adjacent trials as 'not in-order' in kindergarten to 'in-order' by the end of first grade, while another subgroup remained committed to the idea that such trials do not count as being 'in-order' across all three time points. These findings are consistent with the hypothesis that improvement on ONA trials occurs as a qualitative re-classification of in-order adjacent trials from 'not in-order' to 'in-order' (i.e., conceptual shift in ordinal thinking), rather than through incremental improvements.

To probe whether the bimodality observed above was a reflection of a subgroup of children improving on ONA trials specifically, rather than on the ordering task in general, we next divided the sample into those who performed above chance on ONA trials at the end of first grade (T3, $N = 156$), and those who continued to performance below chance on ONA trials at T3 ($N = 228$). We abbreviate these groups as T3+ and T3–, respectively. We then investigated between-group differences in mean performance across all trial-types.

Fig. 4 summarizes mean performance broken down by subgroup (T3+ and T3–) for all conditions at all time-points. First, it is important to note that the T3– group performed above chance in the other three conditions (OA, MA, MNA) at all three time-points (all $ps < 0.001$, all $ds > 0.74$). This indicates that aberrant performance for the T3– group was specific to the ONA condition. This is reinforced by the fact that differences between the two groups in the other conditions were smaller and followed a less consistent pattern (all $ds < 0.50$), relative to differences in the ONA condition at T3 ($d = 4.19$).[4] Indeed, at the critical 3rd time-point (T3, at which group-membership was determined via the ONA condition), the T3– did not perform significantly worse than the T3+ group in any of the other three conditions. If anything, the T3– group showed slightly *better* performance on the two mixed conditions (MA MNA). In other words, it does not appear to be the case that the T3+ group suddenly improved on the ordering task as a whole. Instead, the T3+ group's marked improvement, relative to the T3– group, was specific to the ONA condition, consistent with the notion that the T3+ children had learned to extend the notion of numerical order to sets of non-adjacent numbers (i.e. ordered sets that do not match the count-list). Finally, the T3+ group's improvement on ONA conditions was specific to the period between the 2nd and 3rd time-points, as they showed no advantage over their T3– counterparts on ONA trials at either T1 ($t(383) = 1.33$ $p = .55$, $d = 0.06$) or T2 ($t(383) = -1.27$, $p = .20$, $d = 0.13$). Together, these results are consistent with the notion that a conceptual shift in numerical ordinal understanding occurred in a subset of children – those in the T3+ group – that impacted performance specifically in the ONA condition and was limited to the time-frame between the end of kindergarten and the end of 1st grade.

3.2.2. Research Question 2b – changes in the interrelations amongst trial-types

To complement the qualitative analyses above, we next took a quantitative approach to investigate whether the ability to extend notions of ordinality beyond the count-list develops as a result of a

[4] We recognize that comparing groups on the very dimension by which they were divided is clearly double-dipping. We provide this effect-size here simply to contextualize the relatively smaller differences between the groups in the other three conditions.

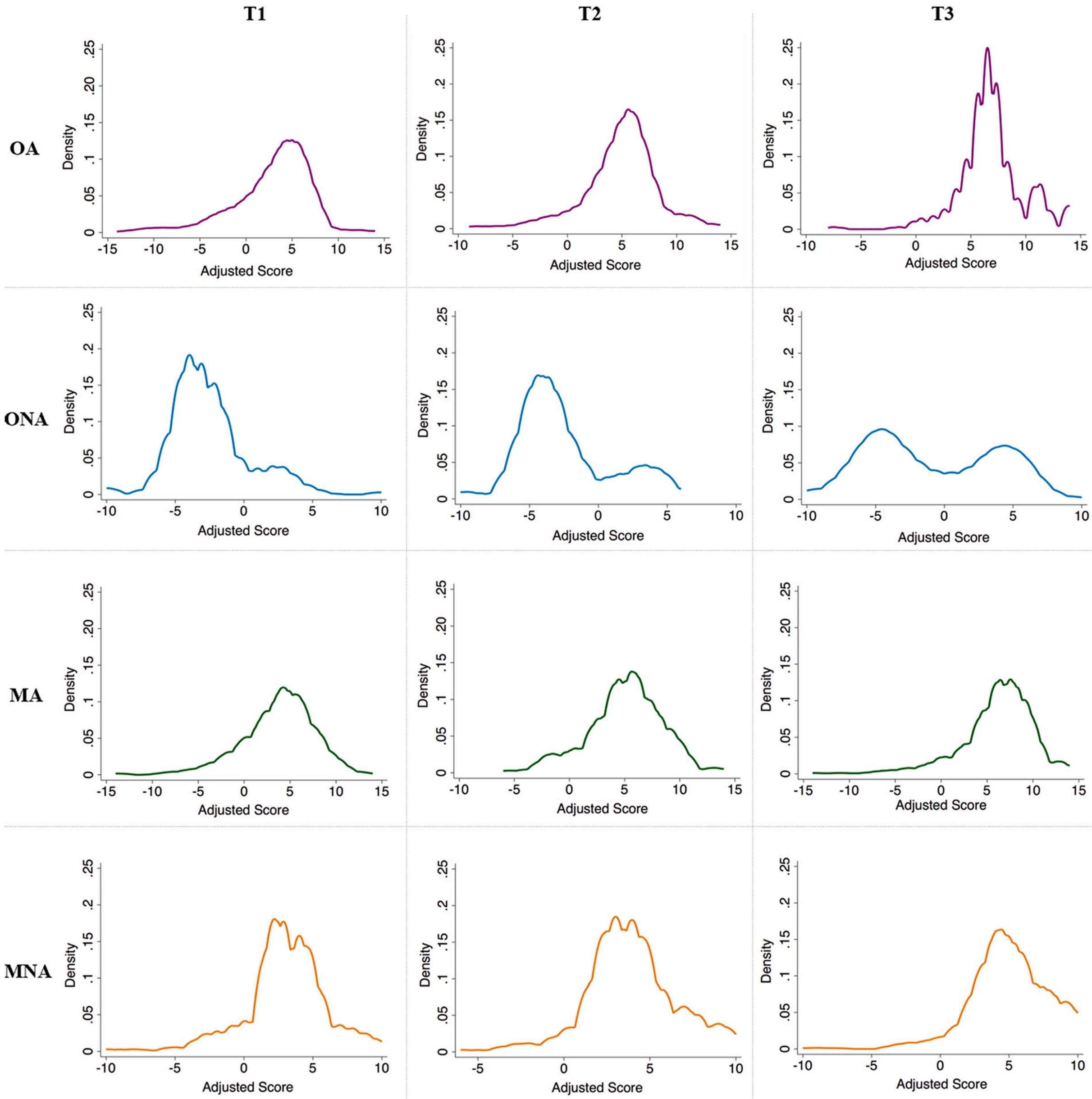


Fig. 3. shows distributions of performance for each trial type across all three time points. The x-axis reflects adjusted scores, while the y-axis reflects sample density. Density distributions were estimated using an Epanechnikov kernel function. OA = in-order adjacent, ONA = in-order non-adjacent, MA = mixed-order adjacent, MNA = mixed-order non-adjacent. T1 = beginning of kindergarten, T2 = end of kindergarten, T3 = end of first grade.

conceptual shift in one's understanding of what it means for a series of numbers to be in-order. Specifically, we examined changes in the interrelations amongst trial-types over time, separately for those who recognized ONA trials as 'in-order' by the end of first grade (T3+) and for those who do not (T3–). If the improved performance for the T3+ group is a reflection of a qualitative re-classification of ONA trials from 'not in-order' to 'in-order', we would expect performance on these trials to relate to performance on the other three trial types in a manner that is substantially different than in the group for which no such re-classification is thought to have occurred (T3–). That is, we expected to see ONA trials relate to the other three trial-types in a fundamentally different way for the T3+ relative to the T3– group. Importantly, this difference should emerge only between time-points 2 and 3, when the putative conceptual shift is thought to have arisen.

Fig. 5 shows correlations between ONA performance and the other three trial types for each subgroup (ONA ~ OA in Fig. 5a, ONA ~ MA in Fig. 5b, and ONA ~ MNA in Fig. 5c). In Fig. 5a, the correlation between ONA and OA trials is moderate at T1 and near zero (and non-significant) at T2 for both groups. However, at T3 the correlation between ONA and OA trials is strongly positive [r (154) = 0.67] for the T3+ group and strongly negative [r (226) = −0.50] for the T3– group. Indeed, comparing ONA ~ OA correlations between groups (using a Fisher z-test) showed no significant difference between groups at T1 [$z = 0.92$, $p = .36$, $d = 0.12$] or T2 [$z = -0.85$, $p = .39$, $d = 0.11$], but there was a highly significant difference at T3 [$z = 13.09$, $p < .001$, $d = 1.73$]. These findings indicate a major departure between groups in terms of how their performance on ONA trials related to OA trials.

An overall similar pattern was seen for the correlations between ONA and MA and ONA and MNA. The correlations were strongly negative for both groups across T1 and T2 and, for the most part, we did not observe any significant differences between groups at either time point ($ps > 0.28$, $ds < 0.14$). However, we did observe a slight between-group difference in the correlation between ONA and MNA at T1 ($z = 2.73$, $p < .01$, $d = 0.36$), in which the correlation was more strongly negative for

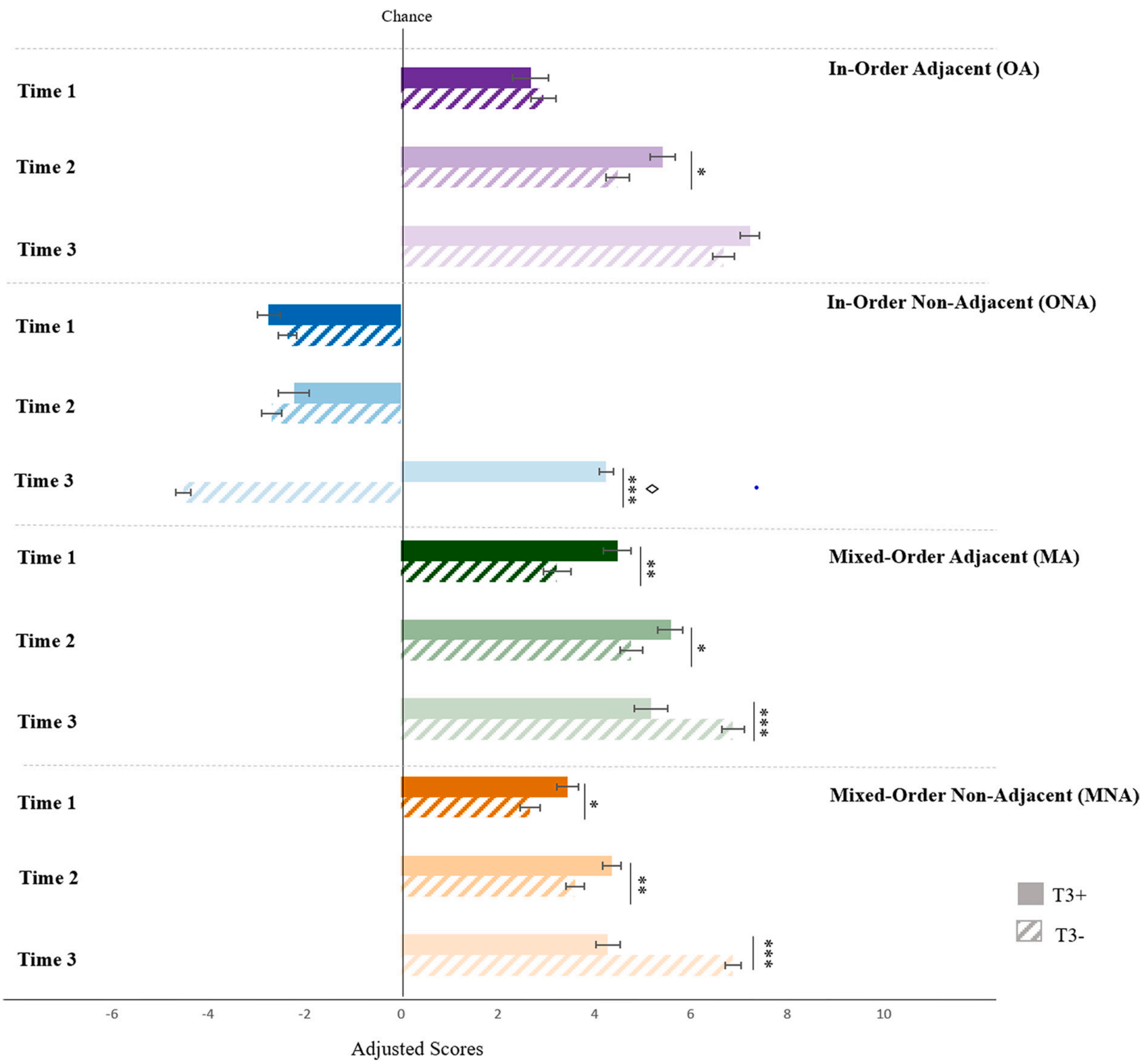


Fig. 4. shows trial-based performance on the ordering task by subgroup. Scores are adjusted so that a score of 0 indicates chance performance. Error bars reflect standard errors. T1 = beginning of kindergarten, T2 = end of kindergarten, T3 = end of first grade. Mean performance for the above chance on ONA trials at T3 subgroup (T3+) is reflected in the solid bars ($N = 156$), while mean performance for the below chance on ONA trials at T3 subgroup (T3–) is reflected in the striped bars ($N = 228$). ⋄Note that based on how the groups were defined, a difference here is necessarily expected. It is presented here not to be of interest in itself, but to provide a reference point for group-wise differences in the other conditions, especially ONA performance at time-points 1 and 2.

the T3– group [r(226) = −0.79] compared to the T3+ group [r(154) = −0.65). Notably, large differences did not emerge until T3 (ONA ~ MA: $z = 9.49, p < .001, d = 1.26$; ONA ~ MNA: $z = 12.07, p < .001, d = 1.60$).

Overall, we observed that belonging to the group of children who reliably identified non-adjacent ordered trials as 'in-order' by the end of first grade, was associated with a significant shift in the interrelations between ONA trials and all other trial-types (shifting from near-zero or negative at T1 and T2, to strongly positive at T3). Crucially, this change was specific to the time-point after which this shift purportedly occurred. These findings suggest that those in the T3+ group experienced an abrupt qualitative change in how they process ONA trials. More broadly, these results support the hypothesis that recognizing numerical order in sequences other than the integer count-list is itself driven by a conceptual shift in numerical ordinal understanding rather than via the type of incremental change we observed for ordinal verification of other types of sequences (OA, MA, MNA).

3.3. Post-hoc analysis: predicting change in non-adjacent ordinal performance

The results above demonstrated that while children at the outset of kindergarten can reliably classify ordinal sequences of numbers, this does not extend to ordered sequences that go beyond the count-list. The capacity to extend conceptions of numerical order began to emerge by the end of 1st grade, but only amongst a subset of children. Moreover, this capacity appeared to qualitatively reorganize how these children approached the ordinal verification task, as evidenced by changing interrelations between trial types. In sum, these results suggest that a deeper investigation into what factors might contribute to this change in children's ordinal understanding may be warranted. Hence in this section, we investigate whether (and which) basic numerical skills predict the change in ONA performance across the time-points that saw the greatest change therein – namely, T2 to T3.

More specifically, we investigated the general and specific contributions of three number comparison tasks that are widely used elsewhere in the literature to measure three foundational aspects of

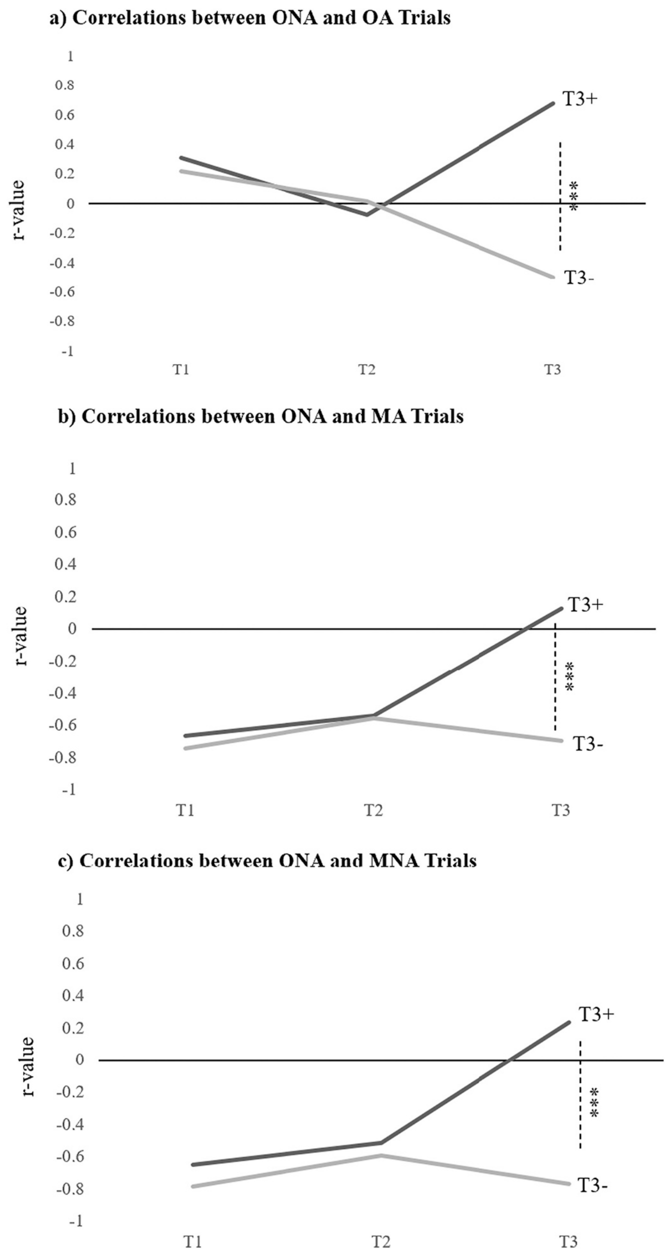


Fig. 5. shows the correlations between performance in the ONA condition and performance in the OA (a), MA (b) and MNA (c) conditions at each of the three time points. T1 = beginning of kindergarten, T2 = end of kindergarten, T3 = end of first grade. The darker line reflects the correlations for the above chance at T3 group (T3+) and the lighter line reflects the correlations for the below chance at T3 group (T3–). ** $p < .01$; *** $p < .001$.

numerical processing: Dot Comparison (DC, nonsymbolic magnitude processing), Numeral Comparison (NC, symbolic magnitude processing), and Mixed-Format Comparison (MC, symbolic-nonsymbolic mapping). We tested whether performance measured at T2 in each of these tasks predicted change in ONA trials between T2 and T3 (predicting ONA_{T3}, controlling for ONA_{T2}). In addition, we tested whether count-list ordinal processing (OA) predicted this change. Importantly, by relying on an individual differences approach, we can account for the fact that we expect some children (T3+ group) but not others (T3– group) to demonstrate improvement on ONA trials; it also allows us to account for variability in improvement within each group, especially in the T3+ group.

From Fig. 6 (blue bars), we see that each of the numerical tasks predicted (positive) change in non-adjacent ordering performance. Thus, if a given child was generally showed stronger basic numerical skills at the end of kindergarten, they were more likely to improve in non-adjacent ordering performance by the end of first grade. However, the red bars in Fig. 6 indicate that only the numerical comparison task (NC) predicted *unique* growth in non-adjacent ordering performance. Hence, it appears that children's understanding of relative cardinal magnitude – especially in symbolic form – contributes to their ability to learn to extend notions of numerical order beyond the count-list.[5]

4. Discussion

Ordinal processing plays a fundamental role in both the representation and manipulation of symbolic numbers (Lyons et al., 2016). As such, it is important to understand how children come to develop a sense of ordinality in the first place. While there is currently a relative dearth of research on how ordinal skills develop, it is reasonable to hypothesize that children's initial perceptions of ordinality are tied to the verbally rehearsed count-list. However, the extent and nature of the count-list's influence on ordinal development remains unknown. The current study examined the role of the count-list in the development of ordinal knowledge through the investigation of two research questions: (1) Do K-1 children struggle to extend the notion of numerical order beyond the count-list, and if so (2) does this extension develop incrementally or manifest as a qualitative re-organization of how children recognize the ordinality of numerical sequences. Overall, we observed that although young children reliably identified adjacent ordered sequences (i.e., those that match the count-list) as being in the correct ascending order, they performed significantly below chance on non-adjacent ordered trials (i.e., those that do not match the count-list but are in the correct order) from the beginning of kindergarten to the end of first grade. This finding suggests that young children's sense of what constitutes numerical order may be limited to sequences that match the count-list.

Further, both qualitative and quantitative analyses supported the conclusion that the ability to extend notions of ordinality beyond the count-list emerged as a conceptual shift in ordinal understanding between the end of kindergarten and the end of first grade, while performance on the other types of ordinal processing developed more incrementally. In sum, we found that the count-list indeed appears to have a profound influence on how children conceptualize numerical order. Specifically, the majority of children found the problem of extending their notion of numerical order beyond the count-list to be highly non-trivial. Further, this extension took place for about half of children sometime between the end of kindergarten and the end of first grade, and appeared to manifest as a qualitative, rather than an incremental, shift in their concept of numerical order. Finally, post hoc analyses revealed that cardinal processing of symbolic numbers may play a unique role in helping children merge concepts of cardinality and ordinality to expand their working understanding of what it means for numbers to be 'in order'.

4.1. *Children struggle to extend their notion of ordinality beyond the count-list*

The finding that young children consistently identified trials that matched the count-list as 'in-order' but struggled with those that did not, is consistent with the well-documented reverse distance effect (RDE). Specifically, prior studies have documented that both adults (Franklin & Jonides, 2009; Lyons & Ansari, 2015; Lyons & Beilock, 2013; Turconi et al., 2006) and children in grades 1–6 (Lyons & Ansari, 2015) are slower and less accurate when identifying non-adjacent ordered sequences compared to adjacent ordered sequences. The current study is

[5] The overall pattern of results remained unchanged after adjusting for clustering within schools using cluster-robust standard errors.

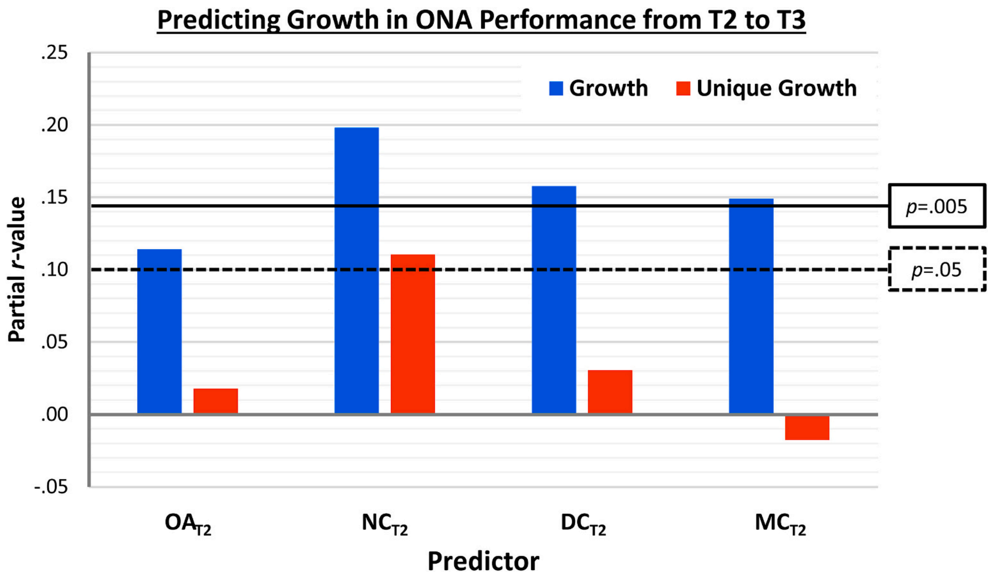


Fig. 6. shows partial correlations predicting growth in non-adjacent ordering (ONA) performance from the end of kindergarten (T2) to the end of first grade (T3). Blue bars show simple growth effects (predicting ONA_{T3}, controlling for ONA_{T2}). Red bars show unique growth effects, controlling for the other three predictors as well. Abbreviations: OA: adjacent ordering, NC: numeral comparison, DC: dot comparison, MC: mixed-format comparison. Dashed and solid black lines indicate $p = .05$ and $p = .005$ significance levels, respectively.

the first to suggest that the RDE obtains in children as early as the beginning of kindergarten.

However, it is important to note that the extent of the RDE observed in the current study differs from that observed in prior studies with older samples. Specifically, although the RDE is typically marked by less efficient processing of non-adjacent ordered sequences compared to adjacent ordered sequences, performance amongst older samples tends to remain well-above chance for both types of sequences. As such, the RDE typically reflects a slight performance cost rather than a fundamental misunderstanding of what it means for a series of numbers to be in-order. However, in the current study, the majority of kindergarten children (and about half of 1st graders) performed significantly *below* chance on non-adjacent ordered trials. This finding suggests that children at this age may not recognize non-adjacent ordered sequences as being in the correct order; in fact, they actively categorize them as *not* in order. The current study is thus the first to provide clear indication that in the early stages of formal schooling (K-1st grade), children have not yet extended their notion of ordinality beyond sequences that directly match the count-list (e.g., 3–4-5).

Until recently, perhaps the most common explanation for the RDE in adults is that it is a by-product of familiarity with rehearsed sequences (such as the count-list). In other words, individuals are faster to recognize adjacent, compared to non-adjacent, sequences of numbers because they are more likely to come across such sequences in their daily lives resulting in more efficient retrieval from long-term memory (Bourassa, 2014; LeFevre & Bisanz, 1986; Lyons & Beilock, 2013; Sella, Sasanguie, & Reynvoet, 2020). However, recent work by Gattas et al. (2021) demonstrated that the RDE is more a reflection of the fact that processing of non-adjacent sequences like 2–4-6 is impeded by the count-list more so than adjacent sequences like 3–4-5 are facilitated by the count-list. In particular, when participants were forced to classify sequences like 2–4-6 as '*not*-in-order' their performance substantially *improved*, indicating an underlying inclination to view such stimuli as poor examples of ordered sequences. Gattas and colleagues surmised this might be a vestige of an earlier developmental struggle to stretch conceptions of numerical order to include non-count-list sequences. The data here are a strong endorsement of that speculation. Furthermore, Gilmore and Batchelor (2021) showed that a subset of older children (roughly 6-9 yrs) continue to perform at or below chance specifically in non-adjacent ordered trials. Together, these results converge to show that overcoming an early tendency to define numerical order as synonymous with the integer count-list is non-trivial, may take several years in some children, and continues to impact ordinal processing even in highly numerate adults.

4.2. Going beyond the count-list

The idea that ordered sets include numerical sequences beyond the count-list is a fundamental numerical concept that allows one to form a rich network of ordinal associations amongst non-adjacent numbers. As such, an important follow-up question is to begin to understand how the expansion of what constitutes the boundaries of numerical order develops. The findings from the current study suggest that the ability to correctly identify non-adjacent ordered sequences as being in the correct order does not develop incrementally, but rather as a qualitative reorganization of what constitutes numerical order. This conclusion is supported by two main findings. For one, we observed that the underlying distribution of performance on non-adjacent ordered trials changed from unimodal (with a peak below the mean) to bimodal (with one peak below and another above the mean) between the end of kindergarten and first grade. If the development of ordinality occurred through incremental change, we would have observed a unimodal distribution of performance at all three time points with a rightward shifting peak – just as was the case for the other three trial types. However, the bimodal distribution at T3 suggests that only a subset of children (about half the sample) transitioned form initially rejecting, to later accepting, the idea that non-adjacent numbers can make up an ordered sequence, with their performance distribution for ONA trials centered above chance (above 0 in Fig. 3) at T3. Meanwhile, the remaining children remained committed to the idea that non-adjacent ordered sequences do not count as being 'in-order', with their distribution centered below 0 at all time-points. Therefore, some children (the T3+ group) 'made the leap' to include ONA trials in their concept of numerical order, while some children (the T3– group) did not. That said, it is still possible that the T3+ group arrived at their expanded sense of order via an incremental process. To address this possibility, we examined whether the T3+ group processed ONA trials differently than their T3– counterparts prior to T3 (1st grade).

First, Fig. 4 shows that the two groups did not differ in overall performance on ONA trials prior to the third time-point. Second, Fig. 5 shows that belonging to the T3+ group was associated with a large shift in the interrelations between ONA trials and the other three trial-types between the end of kindergarten and the end of first grade. Specifically, the relation between ONA and all other trial types shifted from near-zero or negative to positive by the end of first grade. Crucially, this pattern of change was observed specifically for those in the T3+ group.

For those in the T3– group, we observed that the relation between ONA and all other trial-types remained consistently negative or near-zero across all time-points. The abrupt shift in the interrelations between ONA trials and all other trial-types for the ONAT3+ group, but not the T3– group, provides evidence to suggest that the T3+ group underwent a qualitative shift in how they process ONA trials, while the T3– did not. Further, the fact that the interrelations between trial-types for the T3+ group looked more or less identical to their T3– counterparts prior to T3 suggests the relevant developmental change occurred sometime during the 1st grade. Interestingly, the fact that only about half of the sample exhibited such a change even by the end of 1st grade indicates the presence of large individual differences in the timing thereof.

Examining the between-group differences in the *direction* of interrelations amongst trial-types provides further insight into how these two groups differentially perceived ONA trials by the end of first grade. Recall that during the ordinality task children were asked to indicate whether a sequence was 'in-order' or not by responding either 'yes' or 'no'. If children perceive ONA trials as being in the correct numerical order, they should reliably respond 'yes, in-order' to these trials. On the other hand, if children perceive ONA trials as not being in the correct order, they should reliably respond 'no, not in-order' to these trials. Positive correlations between performance on ONA trials and all other trial-types suggest that children are processing in-order *non-adjacent* trials (ONA) similarly to in-order *adjacent* trials (OA; the correct response on these trials is 'yes') and in direct opposition to mixed (MA and MNA) trials (the correct response on these trials is 'no'). On the other hand, negative correlations between performance on ONA trials and all other trial-types suggest that children are processing ONA trials similarly to mixed trials and in direct opposition to in-order adjacent trials. As such, the positive correlations between ONA trials and all other trial-types for the T3+ group suggest that children in this group correctly integrated in-order non-adjacent trials into a more inclusive idea of what constitutes numerical order by the end of first grade. Conversely, the consistent negative correlations between ONA trials and all other trial-types for the T3– group suggest that children in this group continued to incorrectly characterize in-order non-adjacent trials as fundamentally not in-order by the end of first grade.

In sum, only a subgroup of children were able to characterize non-adjacent ordered sequences as being in the correct order even by the end of first grade (less than half here as compared to about a third of children in the older sample from Gilmore & Batchelor, 2021). Belonging to this group was associated with a qualitative shift in how such trials were processed relative to the other trial-types. We argue that these results, taken together, indicate that the extension of ordinality beyond the count-list constitutes a qualitative reorganization of one's sense of what numerical order means.

4.3. Implications and future directions

The findings from the current study provide the first evidence to suggest that over-reliance on the oft-recited count-list may in fact hinder the development of more advanced ordinal processing. In this way, the struggle to acknowledge that sequences such as 2–4-6 are 'in in-order' constitutes another example of how useful concepts at one stage of mathematical learning can prove an impediment to later mathematical learning. We see this as similar to whole-number bias in fraction learning (Mack, 1995; Van Hoof, Verschaffel, & Van Dooren, 2015), in which children often incorrectly select ¼ instead of ½ when asked to determine which is greater, due to overlearned familiarity with integers (i.e., $4 > 2$). In the case of ordering, we propose the count-list may help early acquisition of basic ordinal concepts, but over-reliance on the count-list may lead to an unintended inflexibility – numerical order *is* the count-list – that needs to be partially unlearned. Just as fractions, counter-intuitive as they may be, are an invaluable aspect of mathematics, the notion that *any* set of numbers possesses an ordered permutation is an incredibly powerful mathematical idea. Still more mundanely, simply understanding that sequences can 'skip' numbers is a foundational idea in arithmetic and for developing a rich and flexible network of associations between numerical representations.

To that end, while many studies have documented a strong relation between ordinal processing and arithmetic (Attout & Majerus, 2015; Goffin & Ansari, 2016; Lyons et al., 2014; Lyons & Ansari, 2015; Lyons & Beilock, 2011; Morsanyi et al., 2017; Sasanguie et al., 2017; Sasanguie & Vos, 2018; Vogel et al., 2017), no study has yet investigated whether the ability to extend ordinal principles to include non-adjacent numbers plays a unique role in the development of more complex math skills. However, evidence from dyscalculics suggests that this might be the case (Morsanyi, van Bers, O'Connor, & McCormack, 2018). Specifically, Morsanyi et al., compared the performance of children diagnosed with dyscalculia (i.e., a math learning disability characterized by difficulties with fluency in mathematical operations) and those without math difficulties on a range of tasks including an ordinal verification task. The authors observed that ordering abilities were the best predictor of a dyscalculia diagnosis. This finding contributes to a growing body of literature indicating that deficits in ordinal processing contribute to impaired mathematical processing more broadly (Attout & Majerus, 2015; Attout, Salmon, & Majerus, 2015; De Visscher, Szmalec, Van Der Linden, & Noël, 2015; Kaufmann, Vogel, Starke, Kremser, & Schocke, 2009). Although not acknowledged by the authors, Morsanyi et al. observed that the disparity in ordinal processing between those with and without dyscalculia increased as the distance between the numbers in the sequence being judged increased (Fig. 1, Morsanyi et al., 2018). Specifically, they observed no differences in the ordering abilities of dyscalculics and non-dyscalculics when making judgments about adjacent sequences of numbers (those that match the count-list) but that dyscalculics struggled in comparison to non-dyscalculics when judging the order of three non-adjacent numbers (those that do not match the count-list). Overall, these findings provide some evidence to suggest that the ability to extend one's sense of order beyond the count-list – or perhaps failure to do so – may have important consequences for future acquisition of math skills. Future work may thus prove fruitful if aimed at investigating the consequences of delays in learning this crucial aspect of numbers, as well as the most effective means of helping children make this important conceptual leap.

Further, it is important to consider that the role of ordinal processing in more complex math achievement may be domain general, with both numeric and non-numeric ordering (e.g., alphabet, calendar system, daily events) skills playing an important role (Morsanyi et al., 2017; Morsanyi et al., 2018; Morsanyi, van Bers, O'Connor, & McCormack, 2020; O'Connor, Morsanyi, & McCormack, 2018; Sasanguie et al., 2017; Vos, Sasanguie, Gevers, & Reynvoet, 2017). For example, O'Connor et al. (2018) measured young children's ordering skills using both a symbolic number ordering task, as well as a non-numeric daily events ordering task. The daily events task assessed children's temporal ordering knowledge by asking them to assess the correctness of the order of familiar daily events (e.g., waking up, getting dressed, going to school, eating lunch, eating dinner, going to bed). Overall, the authors observed that children's performance on both the numeric and non-numeric ordering tasks at the beginning of first grade significantly predicted more complex math achievement at later time points. Intriguingly, seminal research on the development of *non-numeric* ordering skills (i.e., the calendar system) suggests that children may also struggle with extending ideas of non-numerical order beyond oft-recited lists (e.g., the verbal month list; Friedman, 1986).

In particular, Friedman observed that fourth graders were significantly faster and more accurate when judging the order of adjacent (e.g., February, March, April) compared to non-adjacent months (e.g., March, May, August). Many fourth graders displayed overt behaviors indicative of reliance on a verbal reciting strategy when asked to determine the order of non-adjacent months (e.g., lip movement, reciting aloud, rhythmic tapping). Friedman interpreted these findings to suggest that familiarity with the verbal month-list may lead to a restricted or inflexible understanding of the temporal month order – similar to our suggestion here that familiarity with the integer count-list may lead to a restricted and inflexible understanding of numerical order. Notably, Friedman observed that the discrepancy in performance on the adjacent and non-adjacent month sequences decreased as children got older, which parallels a reduction in the reverse distance effect with age in numerical order processing (from grades 1–6; see Table 2 in Lyons and Ansari, 2015). Further, Friedman observed that children in tenth grade were much less likely to rely on overt reciting strategies when asked to judge the order of non-adjacent month sequences compared to fourth graders. Friedman therefore concluded that children eventually shift from relying on a verbal list system when making temporal order judgments to a spatially ordered representational system, and that it is this representational shift that eventually allows children to develop a more flexible and mature understanding of the calendar system. Current data do not allow us to infer that a switch to spatially-based processing is what led some children in the current dataset to extend their concept of numerical order beyond the count-list. However, we find the suggestion intriguing and perhaps a ready topic for future research. Overall, we concur with Friedman in that reliance on verbal lists appear to interfere with the development of ordinal knowledge – and this phenomenon may not be limited to numerical processing. Efforts to understand how difficulties or delays in the development of a more flexible understanding of ordinality may interfere with later math achievement should therefore consider the role of both numeric and non-numeric ordering skills (e.g., Morsanyi et al., 2018).

Another factor to consider is that the non-adjacent ordered sequences that children were exposed to in the current study were ‘regular’ sequences, meaning that, although they did not directly match the count-list, there was an equal distance (in this case 2) between each number in the sequence (e.g., 2–4-6, 3–5-7). Through activities such as skip-counting (e.g., counting 2 s), school-age children may have more exposure, and therefore might be more familiar, with ‘regular’ non-adjacent ordered sequences, compared to ‘irregular’ non-adjacent ordered sequences, or those that are not evenly spaced (e.g., 2–5-9; 3–6-8). Given that children in the current study struggled to recognize even ‘regular’ non-adjacent ordered sequences (or those that they potentially have more exposure to) as being in the correct order, one could imagine that this difficulty may be even more pronounced, or last longer, for those that are ‘irregular’. It is possible that learning to extend notions of ordinality to include ‘regular’ non-adjacent ordered sequences first might serve as an intermediary bridge for eventually developing a more complete sense of ordinality that extends to include those that are ‘irregular’ as well. We suggest that future work should include multiple types of non-adjacent ordered trials to investigate this idea.

Finally, the findings from the current study may offer some practical recommendations for educators. While learning the count-list is an integral part of early mathematics education, it is important for teachers to understand that the ability to extend notions of ordinality beyond this list poses a significant challenge to young learners and therefore likely requires explicit scaffolding. Complementing rote-counting activities with skip-counting activities is one way of emphasizing the non-adjacent relations amongst numbers. In fact, skip-counting is included in the Ontario Curriculum as a specific learning expectation in first grade (Ontario Education, 2020), but not kindergarten (Ontario Education, 2016). The introduction of skip-counting in first grade might partially explain why we observed improved performance on non-adjacent ordered trials by the end of first grade, but not across the kindergarten year. Emphasizing the non-adjacent relations amongst numbers earlier on the curriculum may help children develop a more complete sense of ordinality prior to the start of first grade, when children are also expected to learn numerical operations such as addition and subtraction. Entering the first grade classroom with a deeper understanding of the ordinal relations amongst numbers may facilitate the acquisition of these more complex mathematical concepts (Attout & Majerus, 2015; Goffin & Ansari, 2016; Lyons et al., 2014; Lyons & Ansari, 2015; Lyons & Beilock, 2011; Morsanyi et al., 2017; Sasanguie et al., 2017; Sasanguie & Vos, 2018; Vogel et al., 2017). A related point is that Gilmore and Batchelor (2021) found complex counting skills explained the relation between ordinal verification performance and arithmetic performance, whereas Lyons and Ansari (2015) found that simple counting skills did not. That said, neither Gilmore and Batchelor nor Lyons and Ansari explicitly examined skip-counting. Hence, it remains the case that no study has yet investigated the causal influence of skip-counting on the ability to extend ordinal notions beyond the count-list, and what this might mean for other forms of math processing. This remains a potentially fruitful area for future research.

4.4. Why some children go beyond the count-list and others do not

Even if we allow that the introduction of skip-counting in first grade might explain why *average* performance on non-adjacent ordered trials increased between the end of kindergarten and the end of first grade, it does not explain why this only occurred for a subgroup of children. The post hoc analyses indicated that positive change in non-count-list ordinal processing was predicted by stronger basic numeracy skills overall. However, only symbolic number comparison predicted unique change, even controlling for performance on count-list trials. Hence, it seems that some influence outside of current notions of ordinality may be important to expanding those notions. To this end, Spaepen et al. (2018) observed that young children were only able to make ordinal judgments after they had acquired the cardinality principle (i.e., understanding that the last number reached when counting items reflects the size of the whole set). Similarly, using a cross-sectional sample of 88 children (ages 4–7), Knudsen et al. (2015) reported that young children demonstrated an understanding of ordinality only after they had learned to associate number symbols with their underlying qualities. Once children had acquired the meaning of number symbols, cardinal and ordinal understanding developed concurrently. Hence, one possible explanation is that early understanding of symbolic numbers combine with counting procedures to provide an initial platform for completing basic ordinality judgments. However, further development of symbolic cardinal understanding may continue to play a role in helping to expand conceptions of numerical order still further.

It is also important to consider what role environmental factors might play in explaining why some children were quicker to make a conceptual shift in ordinal understanding than others. For one, it is possible that those who were able to identify non-adjacent ordered sequences as being in the correct order by the end of first grade in the current sample had teachers who emphasized the non-adjacent relations amongst numbers (whether through skip-counting or other strategies) to a greater extent than those who did not. However, as mentioned above, more research is needed to understand the extent to which specific teaching practices (such as skip-counting) might support the ability to extend notions of ordinality to include non-adjacent relations amongst numbers. Further, it is possible that those who were able to make this conceptual shift in ordinal understanding were exposed to more numeracy activities at home and therefore had more opportunities to learn about the ordinal relations amongst numbers outside of the classroom. However, although prior research suggests that the home numeracy environment plays an influential role in the development of early numeracy skills in general (e.g., Anders et al., 2012; Elliott & Bachman, 2018; Susperreguy, Di Lonardo Burr, Xu, Douglas, & LeFevre, 2020), more work is needed to understand its role in the development of ordinal understanding specifically.

Finally, in addition to environmental factors, individual differences in non-numerical cognitive skills, such as inhibitory control (i.e., the ability to suppress a pre-potent response; Zelazo, Carlson, & Kesek, 2008) may influence one's ability to extend ordinal principles beyond the count-list. For example, if young children's initial notions of ordinality are rooted in the count-list, then classifying non-adjacent ordered sequences that do not match the count-list as ‘not in-order’ likely becomes a prepotent response that children must overcome in order to respond correctly on these trials. Those with greater inhibitory control should be better able to overcome this prepotent response and therefore may be more likely to belong to the group of children who are able to reliably identify non-adjacent ordered sequences as being in the correct order by the end of first grade. However, the role of inhibitory control in one's ability to make this conceptual shift in ordinal understanding has yet to be empirically investigated. Overall, efforts to understand how various environmental and cognitive factors contribute to one's ability to recognize non-adjacent ordered sequences as being in the correct numerical order can inform why some children appear to struggle with this more than others and may point toward specific strategies for promoting the acquisition of this relatively complex ordinal concept.

4.5. Limitations

While we suggest that the observed underperformance on non-adjacent ordered trials, relative to adjacent ordered trials, stems from a restricted representation of what it means for a sequence of numbers to be 'in-order', others may argue that this pattern of results could stem simply from a misunderstanding of the instructions. For example, it is possible that even if young children equate the term 'in-order' to mean the count-list, they may still have an understanding of ordinal direction for non-adjacent trials (i.e., they recognize that numerical sequences can go up or down whether the numbers are adjacent or non-adjacent in the count-list). In the current study, we only asked about ascending ordered sequences and therefore could not test this hypothesis. However, it is important to note that every effort was made to ensure that participating children did in-fact understand that the term 'in-order' could apply to both adjacent and non-adjacent sequences. Specifically, as mentioned in the Methods section, children completed practice trials that consisted of all trial-types (in-order adjacent, in-order non-adjacent, mixed-order adjacent, mixed-order non-adjacent) and were provided with feedback if they responded incorrectly. Therefore, if a child incorrectly labelled the in-order non-adjacent practice trial as 'not in-order', the experimenter would have corrected them and explained that such trials do in-fact count as being 'in-order'. Importantly, experimenters were instructed that for this task in particular, it is especially important to work through all example problems, providing feedback as needed, to ensure that students understand what it means for numbers to be in the correct order. These instructions were emphasized during both the in-service training day as well as in the task booklet (see Appendix A). It is therefore rather intriguing that despite these efforts to ensure that all children understood that both adjacent and non-adjacent sequences count as being in the correct order, most children continued to reliably classify non-adjacent order sequences as being 'not in-order' (for converging evidence in older children, see also Gilmore & Batchelor, 2021). This suggests that young children may require more extensive scaffolding that goes beyond what was provided in the task instructions in-order to successfully update their representation of what it means for a series of numbers to be in the correct order.

Finally, it is important to acknowledge that the ordinal verification task used in the current study is only one way of measuring numerical ordinal processing. Specifically, in the current study we operationalized ordinal processing as one's ability to identify whether three single-digit numbers were in the correct ascending order or not. While this task is commonly used throughout the literature (e.g., Goffin & Ansari, 2016; LeFevre & Bisanz, 1986; Lyons et al., 2014; Lyons & Ansari, 2015; Lyons & Beilock, 2013), other studies have operationalized ordinal processing using pairs of numbers instead of triplets (e.g., Turconi et al., 2006; Turconi, Jemel, Rossion, & Seron, 2004; Vogel et al., 2017) or by asking participants to arrange a set of numbers in the correct order rather than judging a pre-established sequence (e.g. Knudsen et al., 2015). Future studies should investigate whether the pattern of results observed in the current study holds when ordinality is probed in different ways.

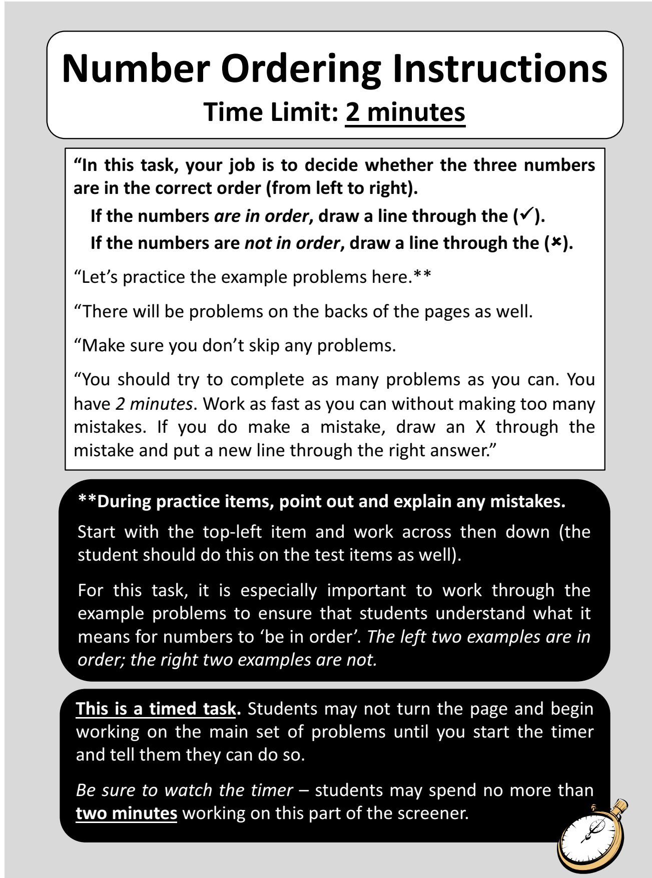
5. Conclusion

Overall, the findings from the current study contribute to a small body of literature that has attempted to uncover the developmental trajectory of ordinal processing. Prior studies suggest that ordinality develops gradually and lags behind the development of cardinal understanding (Colomé & Noël, 2012; Knudsen et al., 2015; Michie, 1984; Spaepen et al., 2018). However, no study has yet considered that different aspects of ordinality may display different developmental trajectories. These findings are the first to suggest that while the ability to recognize sequences that match the count-list as 'in-order' develops gradually, the ability to extend notions of ordinality beyond the count-list to include non-adjacent sequences is non-trivial and reflects a significant developmental hurdle that most children must overcome before they can develop a mature sense of ordinality.

Funding & declarations of interest

This research was supported by funding from the Canadian Institute of Health Research (CIHR), The National Sciences and Engineering Research Council of Canada (NSERC) to Ansari as well as Georgetown University Start-Up Funds to Lyons, and from the Toronto District School Board. There are no declarations of interest to disclose.

Appendix A. Task Instructions



Number Ordering Instructions
Time Limit: 2 minutes

"In this task, your job is to decide whether the three numbers are in the correct order (from left to right).
If the numbers *are in order*, draw a line through the (✓).
If the numbers are *not in order*, draw a line through the (✗).

"Let's practice the example problems here.**

"There will be problems on the backs of the pages as well.

"Make sure you don't skip any problems.

"You should try to complete as many problems as you can. You have *2 minutes*. Work as fast as you can without making too many mistakes. If you do make a mistake, draw an X through the mistake and put a new line through the right answer."

****During practice items, point out and explain any mistakes.**

Start with the top-left item and work across then down (the student should do this on the test items as well).

For this task, it is especially important to work through the example problems to ensure that students understand what it means for numbers to 'be in order'. *The left two examples are in order; the right two examples are not.*

This is a timed task. Students may not turn the page and begin working on the main set of problems until you start the timer and tell them they can do so.

Be sure to watch the timer – students may spend no more than **two minutes** working on this part of the screener.

Fig. A.1. The exact directions and script provided to experimenters.

Appendix B. Trial-List

Table B.1
Ordering trials in the order that they appear in the task booklet.

Trial #	Sequence					Type
1	2	–	3	–	4	In-Order Adjacent (OA)
2	2	–	1	–	3	Mixed-Order Adjacent (MA)
3	7	–	9	–	8	Mixed-Order Adjacent (MA)
4	3	–	4	–	5	In-Order Adjacent (OA)
5	1	–	2	–	3	In-Order Adjacent (OA)
6	7	–	5	–	9	Mixed-Order Non-Adjacent (MNA)
7	5	–	6	–	7	In-Order Adjacent (OA)
8	6	–	2	–	4	Mixed-Order Non-Adjacent (MNA)
9	6	–	4	–	5	Mixed-Order Adjacent (MA)
10	1	–	3	–	5	In-Order Non-Adjacent (ONA)
11	6	–	7	–	8	In-Order Adjacent (OA)
12	3	–	5	–	7	In-Order Non-Adjacent (ONA)
13	6	–	8	–	7	Mixed-Order Adjacent (MA)
14	2	–	4	–	3	Mixed-Order Adjacent (MA)
15	5	–	7	–	9	In-Order Non-Adjacent (ONA)
16	8	–	4	–	6	Mixed-Order Non-Adjacent (MNA)
17	4	–	5	–	6	In-Order Adjacent (OA)
18	4	–	6	–	8	In-Order Non-Adjacent (ONA)
19	5	–	3	–	7	Mixed-Order Non-Adjacent (MNA)
20	4	–	5	–	3	Mixed-Order Adjacent (MA)
21	6	–	7	–	5	Mixed-Order Adjacent (MA)
22	7	–	8	–	9	In-Order Adjacent (OA)
23	3	–	5	–	1	Mixed-Order Non-Adjacent (MNA)
24	2	–	4	–	6	In-Order Non-Adjacent (ONA)
25	4	–	6	–	5	Mixed-Order Adjacent (MA)
26	1	–	3	–	2	Mixed-Order Adjacent (MA)
27	4	–	2	–	6	Mixed-Order Non-Adjacent (MNA)
28	2	–	3	–	4	In-Order Adjacent (OA)
29	7	–	3	–	5	Mixed-Order Non-Adjacent (MNA)

(continued on next page)

Table B.1 (*continued*)

Trial #	Sequence					Type
30	4	–	6	–	8	In-Order Non-Adjacent (ONA)
31	8	–	7	–	9	Mixed-Order Adjacent (MA)
32	6	–	7	–	8	In-Order Adjacent (OA)
33	3	–	4	–	5	In-Order Adjacent (OA)
34	5	–	6	–	7	In-Order Adjacent (OA)
35	3	–	1	–	5	Mixed-Order Non-Adjacent (MNA)
36	5	–	9	–	7	Mixed-Order Non-Adjacent (MNA)
37	2	–	4	–	6	In-Order Non-Adjacent (ONA)
38	4	–	5	–	6	In-Order Adjacent (OA)
39	5	–	3	–	4	Mixed-Order Adjacent (MA)
40	7	–	5	–	6	Mixed-Order Adjacent (MA)
41	3	–	5	–	7	In-Order Non-Adjacent (ONA)
42	7	–	8	–	6	Mixed-Order Adjacent (MA)
43	6	–	8	–	4	Mixed-Order Non-Adjacent (MNA)
44	1	–	2	–	3	In-Order Adjacent (OA)
45	3	–	4	–	2	Mixed-Order Adjacent (MA)
46	5	–	7	–	9	In-Order Non-Adjacent (ONA)
47	7	–	8	–	9	In-Order Adjacent (OA)
48	1	–	3	–	5	In-Order Non-Adjacent (ONA)

References

Anders, Y., Rossbach, H. G., Weinert, S., Ebert, S., Kuger, S., Lehrl, S., & Von Maurice, J. (2012). Home and preschool learning environments and their relations to the development of early numeracy skills. *Early Childhood Research Quarterly, 27*(2), 231–244. https://doi.org/10.1016/j.ecresq.2011.08.003

Attout, L., & Majerus, S. (2015). Working memory deficits in developmental dyscalculia: The importance of serial order. *Child Neuropsychology, 21*(4), 432–450. https://doi.org/10.1080/09297049.2014.922170

Attout, L., Salmon, E., & Majerus, S. (2015). Working memory for serial order is dysfunctional in adults with a history of developmental dyscalculia: Evidence from behavioral and neuroimaging data. *Developmental Neuropsychology, 40*(4), 230–247. https://doi.org/10.1080/87565641.2015.1036993

Bourassa, A. (2014). *Numerical sequence recognition: is familiarity or ordinality the primary factor in performance?*. Masters Thesis. Carleton University. https://doi.org/10.22215/etd/2014-10567.

Colomé, À., & Noël, M.-P. (2012). One first? Acquisition of the cardinal and ordinal uses of numbers in preschoolers. *Journal of Experimental Child Psychology, 113*(2), 233–247. https://doi.org/10.1016/j.jecp.2012.03.005

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

Elliott, L., & Bachman, H. J. (2018). How do parents foster young children's math skills? *Child Development Perspectives, 12*(1), 16–21. https://doi.org/10.1111/cdep.12249

Franklin, M. S., & Jonides, J. (2009). Order and magnitude share a common representation in parietal cortex. *Journal of Cognitive Neuroscience, 21*(11), 2114–2120. https://doi.org/10.1162/jocn.2008.21181

Friedman, W. J. (1986). The development of children's knowledge of temporal structure. *Child Development*, 1386–1400. https://doi.org/10.2307/1130418

Gattas, S. U., Bugden, S., & Lyons, I. M. (2021). Rules of order: Evidence for a novel influence on ordinal processing of numbers. *Journal of Experimental Psychology: General*. https://doi.org/10.1037/xge0001022. Advance online publication.

Gebuis, T., & Reynvoet, B. (2012). The interplay between nonsymbolic number and its continuous visual properties. *Journal of Experimental Psychology: General, 141*(4), 642. https://doi.org/10.1037/a0026218

Gilmore, C., & Batchelor, S. (2021). Verbal count sequence knowledge underpins numeral order processing in children. *Acta Psychologica, 216*, 103294. https://doi.org/10.1016/j.actpsy.2021.103294

Goffin, C., & Ansari, D. (2016). Beyond magnitude: Judging ordinality of symbolic number is unrelated to magnitude comparison and independently relates to individual differences in arithmetic. *Cognition, 150*, 68–76. https://doi.org/10.1016/j.cognition.2016.01.018

Hutchison, J. E., Ansari, D., Zheng, S., Jesus, S. D., & Lyons, I. M. (2020). The relation between subitizable symbolic and non-symbolic number processing over the kindergarten school year. *Developmental Science, 23*(2), Article e12884. https://doi.org/10.1111/desc.12884

Jordan, N. C., Kaplan, D., Oláh, L. N., & Locuniak, M. N. (2006). Number sense growth in kindergarten: A longitudinal investigation of children at risk for mathematics difficulties. *Child Development, 77*(1), 153–175. https://doi.org/10.1111/j.1467-8624.2006.00862.x

Kaufmann, L., Vogel, S. E., Starke, M., Kremser, C., & Schocke, M. (2009). Numerical and non-numerical ordinality processing in children with and without developmental dyscalculia: Evidence from fMRI. *Cognitive Development, 24*(4), 486–494. https://doi.org/10.1016/j.cogdev.2009.09.001

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

LeFevre, J.-A., & Bisanz, J. (1986). A cognitive analysis of number-series problems: Sources of individual differences in performance. *Memory & Cognition, 14*(4), 287–298. https://doi.org/10.3758/BF03202506

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., & Beilock, S. L. (2009). Beyond quantity: Individual differences in working memory and the ordinal understanding of numerical symbols. *Cognition, 113*(2), 189–204. https://doi.org/10.1016/j.cognition.2009.08.003

Lyons, I. M., & Beilock, S. L. (2011). Numerical ordering ability mediates the relation between number-sense and arithmetic competence. *Cognition, 121*(2), 256–261. https://doi.org/10.1016/j.cognition.2011.07.009

Lyons, I. M., & Beilock, S. L. (2013). Ordinality and the nature of symbolic numbers. *The Journal of Neuroscience, 33*(43), 17052–17061. https://doi.org/10.1523/JNEUROSCI.1775-13.2013

Lyons, I. M., Bugden, S., Zheng, S., De Jesus, S., & Ansari, D. (2018). Symbolic number skills predict growth in nonsymbolic number skills in kindergarteners. *Developmental Psychology, 54*(3), 440–457. https://doi.org/10.1037/dev0000445

Lyons, I. M., Price, G. R., Vaessen, A., Blomert, L., & Ansari, D. (2014). Numerical predictors of arithmetic success in grades 1–6. *Developmental Science, 17*(5), 714–726. https://doi.org/10.1111/desc.12152

Lyons, I. M., Vogel, S. E., & Ansari, D. (2016). On the ordinality of numbers. In *, Vol. 227. Progress in Brain Research* (pp. 187–221). Elsevier. https://doi.org/10.1016/bs.pbr.2016.04.010.

Mack, N. K. (1995). Critical ideas, informal knowledge, and understanding fractions. In *Providing a foundation for teaching mathematics in the middle grades* (pp. 67–84).

Merkley, R. (2015). *Beyond number sense: Contributions of domain-general processes to the development of numeracy in early childhood.* PhD thesis. Kauf: University of Oxford.

Merkley, R., Shimi, A., & Scerif, G. (2016). Electrophysiological markers of newly acquired symbolic numerical representations: The role of magnitude and ordinal information. *ZDM, 48*(3), 279–289. https://doi.org/10.1007/s11858-015-0751-y

Michie, S. (1984). Number understanding in preschool children. *British Journal of Educational Psychology, 54*(3), 245–253. https://doi.org/10.1111/j.2044-8279.1984.tb02588.x

Morsanyi, K., O'Mahony, E., & McCormack, T. (2017). 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. *Quarterly Journal of Experimental Psychology, 70*(12), 2497–2517. https://doi.org/10.1080/17470218.2016.1246577

Morsanyi, K., van Bers, B. M., O'Connor, P. A., & McCormack, T. (2018). Developmental dyscalculia is characterized by order processing deficits: Evidence from numerical and non-numerical ordering tasks. *Developmental Neuropsychology, 43*(7), 595–621.

Morsanyi, K., van Bers, B. M., O'Connor, P. A., & McCormack, T. (2020). The role of numerical and non-numerical ordering abilities in mathematics and reading in middle childhood. *Contemporary Educational Psychology, 62*, 101895. https://doi.org/10.1016/j.cedpsych.2020.101895

Nosworthy, N., Bugden, S., Archibald, L., Evans, B., & Ansari, D. (2013). A two-minute paper-and-pencil test of symbolic and nonsymbolic numerical magnitude processing explains variability in primary school children's arithmetic competence. *PLoS One, 8* (7), Article e67918. https://doi.org/10.1371/journal.pone.0067918

Núñez, R. E. (2017). Is there really an evolved capacity for number? *Trends in Cognitive Sciences, 21*(6), 409–424. https://doi.org/10.1016/j.tics.2017.03.005

O'Connor, P. A., Morsanyi, K., & McCormack, T. (2018). Young children's non-numerical ordering ability at the start of formal education longitudinally predicts their symbolic number skills and academic achievement in maths. *Developmental Science, 21*(5), Article e12645. https://doi.org/10.1111/desc.12645

Ontario Education. (2016). *The Ontario curriculum, kindergarten.* Ontario: Ministry of Education.

Ontario Education. (2020). *The Ontario curriculum, grades 1–8: Mathematics*. Ontario: Ministry of Education.

Piazza, M., Pica, P., Izard, V., Spelke, E. S., & Dehaene, S. (2013). Education enhances the acuity of the nonverbal approximate number system. *Psychological Science, 24*(6), 1037–1043. https://doi.org/10.1177/0956797612464057

Rowley, G. L., & Traub, R. E. (1977). Education enhances the acuity of the nonverbal approximate number system. *Psychological Science, 14*(1), 15–22. https://www.jstor.org/stable/1433851.

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

Sasanguie, D., & Vos, H. (2018). About why there is a shift from cardinal to ordinal processing in the association with arithmetic between first and second grade. *Developmental Science, 21*(5), Article e12653. https://doi.org/10.1111/desc.12653

Sella, F., Sasanguie, D., & Reynvoet, B. (2020). Judging the order of numbers relies on familiarity rather than activating the mental number line. *Acta Psychologica, 204*, 103014. https://doi.org/10.1016/j.actpsy.2020.103014

Spaepen, E., Gunderson, E. A., Gibson, D., Goldin-Meadow, S., & Levine, S. C. (2018). Meaning before order: Cardinal principle knowledge predicts improvement in understanding the successor principle and exact ordering. *Cognition, 180*, 59–81. https://doi.org/10.1016/j.cognition.2018.06.012

Susperreguy, M. I., Di Lonardo Burr, S., Xu, C., Douglas, H., & LeFevre, J. A. (2020). Children's home numeracy environment predicts growth of their early mathematical skills in kindergarten. *Child Development.*. https://doi.org/10.1111/cdev.13353

Turconi, E., Campbell, J. I. D., & Seron, X. (2006). Numerical order and quantity processing in number comparison. *Cognition, 98*(3), 273–285. https://doi.org/10.1016/j.cognition.2004.12.002

Turconi, E., Jemel, B., Rossion, B., & Seron, X. (2004). Electrophysiological evidence for differential processing of numerical quantity and order in humans. *Cognitive Brain Research, 21*(1), 22–38. https://doi.org/10.1016/j.cogbrainres.2004.05.003

Van Hoof, J., Verschaffel, L., & Van Dooren, W. (2015). Inappropriately applying natural number properties in rational number tasks: Characterizing the development of the natural number bias through primary and secondary education. *Educational Studies in Mathematics, 90*(1), 39–56.

Vogel, S. E., Haigh, T., Sommerauer, G., Spindler, M., Brunner, C., Lyons, I. M., & Grabner, R. H. (2017). Processing the order of symbolic numbers: A reliable and unique predictor of arithmetic fluency. *Journal of Numerical Cognition, 3*(2), 288–308. https://doi.org/10.5964/jnc.v3i2.55

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, 29*(4), 469–482. https://doi.org/10.1080/20445911.2017.1282490

Wynn, K. (1992). 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

Zelazo, P. D., Carlson, S. M., & Kesek, A. (2008). The development of executive function in childhood. In C. A. Nelson, & M. Luciana (Eds.), *Developmental cognitive neuroscience. Handbook of developmental cognitive neuroscience* (pp. 553–574). MIT Press.