



**QUEEN'S  
UNIVERSITY  
BELFAST**

## **Developmental dyscalculia is characterised by order processing deficits: Evidence from numerical and non-numerical ordering tasks**

Morsanyi, K., van Bers, B., O'Connor, P., & McCormack, T. (2018). Developmental dyscalculia is characterised by order processing deficits: Evidence from numerical and non-numerical ordering tasks. *Developmental Neuropsychology*. <https://doi.org/10.1080/87565641.2018.1502294>

**Published in:**  
Developmental Neuropsychology

**Document Version:**  
Peer reviewed version

**Queen's University Belfast - Research Portal:**  
[Link to publication record in Queen's University Belfast Research Portal](#)

**Publisher rights**  
© 2018 Taylor & Francis. This work is made available online in accordance with the publisher's policies. Please refer to any applicable terms of use of the publisher.

**General rights**  
Copyright for the publications made accessible via the Queen's University Belfast Research Portal is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

**Take down policy**  
The Research Portal is Queen's institutional repository that provides access to Queen's research output. Every effort has been made to ensure that content in the Research Portal does not infringe any person's rights, or applicable UK laws. If you discover content in the Research Portal that you believe breaches copyright or violates any law, please contact [openaccess@qub.ac.uk](mailto:openaccess@qub.ac.uk).

### Abstract

This study tested the hypothesis that individuals with dyscalculia have an order processing deficit. The ordering measures included both numerical and non-numerical ordering tasks, and ordering of both familiar and novel sequences was assessed. Magnitude processing/estimation tasks, and measures of inhibition skills were also administered. The participants were 20 children with developmental dyscalculia, and 20 children without maths difficulties. The two groups were closely matched on age, gender, socio-economic status, educational experiences, IQ and reading ability. The findings revealed differences between the groups in both ordering and magnitude processing skills. Nevertheless, diagnostic status was best predicted by order processing abilities.

Keywords: dyscalculia; magnitude processing; mathematics difficulties; number line; order processing.

Developmental dyscalculia (DD) is a specific impairment of mathematical ability, which may affect 3.5 – 6.5% of the population (e.g., Butterworth, 2005; Kaufmann & von Aster, 2012; von Aster & Shalev, 2007). Individuals with DD are characterized by moderate to extreme difficulties in fluent numerical computations in the absence of sensory difficulties, low IQ, or educational deprivation (Butterworth, 2005). Different theories regarding the causes of DD have been proposed. A cognitive neuroscience theory that has dominated research into DD for several years assumes that the specific difficulties in mathematics originate in the impairment of a specialised magnitude representation system, the approximate number system (ANS; Piazza, Facoetti, Trussardi, Berteletti, Conte, Lucangeli, Dehaene & Zorzi, 2010) or ‘number module’ (Landerl, Bevan & Butterworth, 2004). Others have suggested that DD results from impaired connections between these magnitude representations and numerical symbols (Iuculano, Tang, Hall & Butterworth, 2008; Rouselle & Noël, 2007; De Smedt & Gilmore, 2011).

Beside theories of DD that are related to magnitude representation, behavioural research has proposed other theories that focus on the role of more general cognitive resources in mathematics. For example, there is evidence that deficits in verbal and visual working memory (Bull & Scerif, 2001; Geary, 2004, 2011; Hitch & McAuley, 1991; Mammarella, Hill, Devine, Caviola, & Szűcs, 2015; Passolunghi & Siegel, 2001, 2004; Szűcs, Devine, Soltész, Nobes & Gabriel, 2013a; Swanson, 2011), inhibitory function (Blair & Razza, 2007; Bull & Scerif, 2001; Espy, McDiarmid, Cwik, Stalets, Hamby & Senn, 2004; Szűcs et al., 2013a; Swanson, 2011) and attentional function (Ashkenazi, Rubinsten, & Henik, 2009; Hannula, Lepola & Lehtinen, 2010; Szűcs et al., 2013a; Swanson, 2011) may be linked to DD.

More recently, the relation between order processing and mathematics ability has attracted much attention. According to Rubinsten and Sury (2011), numerical cognition might

depend on two core systems, one for representing magnitudes (i.e., cardinality) and one for representing ordinal information. These authors have proposed that it might be the system responsible for order processing (rather than the magnitude system) that is impaired in DD. Indeed, in recent years, an increasing number of studies have investigated the role of order processing in maths, demonstrating its role in both typical mathematical development (e.g., Attout, Noël & Majerus, 2015; Lyons & Ansari, 2015; Lyons, Price, Vaessen, Blomert & Ansari, 2014; Lyons, Vogel & Ansari, 2016; O'Connor, Morsanyi & McCormack, 2018; Vogel, Remark & Ansari, 2015) and in mature mathematics skills (e.g., Goffin & Ansari, 2016; Lyons & Beilock, 2011). [Most of these studies have focussed on number ordering ability, which is typically measured using a task where three one-digit numbers are presented \(e.g., 6 4 7\), and participants have to decide if these number are in the correct order with regard to their position in the count list.](#) Some studies with adults (Morsanyi, O'Mahony & McCormack, 2017; Sasanguie, Lyons, De Smedt & Reynvoet, 2017; Vos, Sasanguie, Gevers & Reynvoet, 2017) not only showed a close link between number ordering and arithmetic skills, but they also found strong relations between non-numerical ordering abilities and maths. [One particular task that has been used in this literature is the month ordering task, where participants have to judge if three months \(e.g., May June August\) are presented in the correct order with regard to their order within a calendar year.](#) ~~Additionally,~~ [In addition to these lab-based studies,](#) clinical observations of dyscalculia also describe characteristic difficulties with everyday activities that require ordering skills (e.g., recalling the order of past events, following sequential instructions, etc., National Center for Learning Disabilities, 2007).

Nevertheless, apart from Rubinsten and Sury (2011), so far only a handful of studies investigated the role of order processing abilities in DD (Attout, & Majerus, 2014; Attout, Salmon & Majerus, 2015; Kaufmann, Vogel, Starke, Kremser & Schocke, 2011). These

Formatted: Indent: First line: 1.25 cm

studies provided initial evidence that not only numerical, but also non-numerical ordering skills are impaired in dyscalculia. However, ~~as discussed below~~, these studies have used a restricted range of ordering tasks (including the number ordering task, as well as a serial order memory task, which requires participants to recall the order of items within a list, where the requirement to remember the items is minimized, as items are presented to the participants during the recall phase – see more details below). Additionally, Kaufmann et al. (2011) only found differences in brain activation patterns between their dyscalculic and non-dyscalculic samples, but they did not find behavioural differences in ordering abilities.

An interesting study by De Visscher, Szmalec, Van der Linden and Noel (2015) investigated serial learning ability in dyscalculia, using the Hebb learning paradigm, where participants were repeatedly presented with lists consisting of nine syllables until they learnt to recall the sequences in the correct order. De Visscher et al (2015) found that participants with dyscalculia showed a reduced ability to learn and retain the sequences. Additionally, a second group of participants, with impaired arithmetic fluency (but no global maths deficit) showed increased sensitivity to interference when the sequences that they had to learn included overlapping items. Given that the sequences used in this study included non-numerical items, these results suggest that dyscalculic participants show a general deficit in serial order learning (which is likely to also affect the learning of numerical sequences, although this was not investigated in the study).

Formatted: Indent: First line: 1.25 cm

#### *The current study*

The participants in our study were 9- and 10-year-old children with DD, and a matched control group. We decided to focus on this age group because several previous

studies on DD included child participants of a similar age, and dyscalculia can be diagnosed with confidence at this age. Specifically, we only recruited children who had shown sustained difficulties in mathematics over several school years. We did this on the basis of the schools' records of children's standardized test scores in mathematics in previous school years, as well as on the basis of their scores on standardized maths tests that were administered by us. Additionally, in line with most previous studies on DD, we selected children with specific difficulties in mathematics ([i.e., not a general intellectual deficit, which would correspond to a different diagnostic category in the DSM-5](#)).

#### *Order processing*

Given that order processing in dyscalculia is a relatively neglected issue, the main aim of the study was to conduct a systematic investigation into ordering skills in DD. Specifically, we tested ordering abilities both in the domain of numbers and in non-numerical domains, such as the domain of time, and everyday activities with an ordering component. Another question that we wanted to address was whether the ordering deficits (if present) affect ordering skills related to familiar sequences (e.g., numbers, familiar everyday events, and the events of the calendar year), as well as short novel sequences (e.g., lists of unrelated words or spatial locations). This way we could assess both the extent and the nature of the order processing deficits in DD.

Although it might be thought that ordering of familiar and unfamiliar sequences rely on similar processes, there are good reasons to assume that these are distinct skills. Tasks that assess memory for short unfamiliar sequences of items assess a key aspect of short-term or working memory ([and is subject to effects related to the limited capacity of working memory resources](#)), whereas tasks that involve making judgments about the order of familiar sequences (either numerical or non-numerical) rely on retrieving and using ordinal

information stored in long-term memory (cf., Attout et al., 2015). Although it can be expected that the two skills are related (for example, it is necessary to temporarily retain sequences in working memory in order to transfer them into long-term memory storage), it is possible that a person who struggles with the temporary storage of arbitrary sequences, after sufficient practice, is able to store these sequence in long-term memory. -Consistent with the idea that tasks assessing order WM and tasks that require long-term memory for familiar sequences tap at least partially independent ~~different~~ skills, O'Connor et al. (2018) found that the two tasks showed a moderate relationship, and that performance on the latter type of task was predictive of maths skills over and above the former. In this study, we looked at both types of abilities.

Regarding the choice of ordering tasks, the two most relevant previous studies for our purposes, which investigated ordering abilities in 10-year-old ~~children with~~ DD, were the studies by Attout and Majerus (2014) and Kaufmann et al. (2011). Both studies used a number ordering task, which required that participants judge the correctness of the order of three numerals. Whereas Kaufmann et al. (2011) only found differences between dyscalculic and non-dyscalculic children in their brain activation patterns, Attout and Majerus (2014) reported a difference in reaction times, with longer RTs in the DD group. Additionally, Attout and Majerus (2014) also used an order working memory task (the animal race task), which involved recalling lists of animal names in the correct order in a task that placed minimal demands on item memory. Attout and Majerus (2014) found impaired order memory in children with DD. Subsequently, they replicated this finding with adults with a history of dyscalculia (Attout et al., 2015). In order to investigate order processing skills in DD further, in the current study, we used both the order memory and the number ordering tasks. Given reports that spatial working memory is impaired in dyscalculia (e.g., Mammarella et al.,

2015; Szűcs et al., 2013a), and because the order memory task only included verbal materials, we also administered a visual-spatial working memory task that also required order recall.

The novelty of our study was the addition of two further measures of non-numerical order processing. Although Attout and Majerus (2014) assessed non-numerical order processing, their task only assessed the ability to hold ordered sequences temporarily in working memory. Their numerical ordering task (judging whether three numbers are in the correct order), performance on which was also impaired in DD, draws on long-term memory representations of a familiar sequence. What is not known is whether the impairments in DD extend to tasks involving long-term memory representations of non-numerical order. Such impairments would indicate that DD is characterized by more generalized difficulties in representing and using order information. Such generalized difficulties would be consistent with recent findings that non-numerical order processing is strongly related to maths skills in typical adults (Morsanyi et al., 2017; Vos et al., 2017).

Indeed, clinical reports of dyscalculia describe difficulties with everyday activities with an ordering requirement. The link between mathematics skills and everyday ordering abilities was supported by a recent study (O'Connor et al., 2018), which was conducted with young children at the start of their formal education. Parents were asked to complete a short questionnaire that included items, such as: "My son/daughter would be able to recall the order of typical daily events". The parents' ratings of their children's everyday ordering abilities strongly predicted their formal mathematics skills at the end of their first, as well as their second school year, even after controlling for the effect of the socio-economic status and IQ of the children. In the current study, we used this questionnaire with some modifications to make the questions age-appropriate.

The same study by O'Connor et al. (2018) found that young children's performance on a non-numerical ordering task at the start of the first school year also strongly predicted

their formal maths skills at the end of the first and second year of school. The task that they used was designed to be similar to the number ordering task. However, instead of a triad of numbers, pictures of three familiar daily events were presented (e.g., having breakfast, going to bed at night, going to school), and children were asked to judge whether the order was correct or incorrect. A strong link between ordering in the time domain ([i.e., ordering of the months of the calendar year](#)) and maths skills was also reported by two recent studies with adults (Morsanyi et al., 2017; Vos et al., 2017)<sup>1</sup>. For this reason, we also administered a task which required the correct ordering of familiar events in time. O'Connor et al. (2018) used pictures of familiar daily events, and the adult studies used written month names. In the current study, we relied on pictorial presentations, so that reading ability did not affect ordering performance. However, instead of pictures of daily events, we presented the children with pictures of familiar events of the calendar year. Whereas even 4-year-olds can make correct judgments about the order of typical daily events, children gradually develop representations of the relative order of recurring events of the calendar year during middle-childhood (Friedman, 1990; 2000; 2002). For this reason, we considered the annual event ordering task as more age-appropriate than a daily event ordering task.

#### *Magnitude processing and estimation*

In addition to ordering skills, we investigated magnitude processing/estimation skills that are traditionally considered to be impaired in dyscalculia. [We thought that it was important to investigate these accounts side by side, in order to see which tasks discriminate the best between dyscalculic and non-dyscalculic children, and also to see if ordering and magnitude processing skills are related.](#) On the basis of past research, three tasks seemed particularly relevant: the number comparison task (which measures symbolic comparison

---

<sup>1</sup> Related to this point, some studies (e.g., Skagerlund & Traff, 2014; Vicario, Rappo, Pepi, Pavan & Martino, 2012) have found impaired time estimation abilities in dyscalculia. However, this skill was not investigated in the current study.

skills), the dot comparison task (which measures non-symbolic magnitude comparison abilities), and the number line task (which measures the ability to translate between symbolic and non-symbolic representations of magnitudes, and is also assumed to include an ordering component).

Although there is evidence to suggest that children with DD show impairments on both the number comparison task and the dot comparison task, some researchers have argued that non-symbolic magnitude processing might be intact in DD, whereas the processing of symbolic magnitudes is impaired (Iuculano et al., 2008; Rouselle & Noël, 2007; De Smedt & Gilmore, 2011). Indeed, both the behavioural and neuroscience evidence regarding differences on the dot comparison task between DD and non-DD participants is conflicting (see Szűcs et al, 2013 for a review and discussion), meaning further investigation is valuable. We thus included a dot comparison task, modelled ~~our task~~ on that used in Price, Holloway, Räsänen, Vesterinen and Ansari's (2007) study which has provided the strongest evidence for group differences so far, showing both behavioural and a neural differences between the DD and control participants.

A final task related to estimation/magnitude processing skills that we used was the number line task, which is a strong predictor of mathematics skills in the first school years (e.g., Booth & Siegler, 2006, 2008; De Smedt, Verschaffel, & Ghesquiere, 2009; Lyons et al., 2014). Additionally, von Aster and Shalev (2007) proposed that the core deficit characterising DD was an impaired ability to represent and manipulate numerical magnitudes nonverbally on an internal number line. In line with this claim, Skagerlund and Traff (2016) found greater estimation errors in a number line task in dyscalculic than in typically developed children. This task is also very relevant for our study, because it is considered to combine the need for estimation skills with an understanding of ordinality (cf., von Aster & Shalev, 2007; Kaufmann & von Aster, 2012).

### *Inhibition*

Inhibition skills are important for mathematics (e.g., Bull & Scerif, 2001; Blair & Razza, 2007), and Szűcs et al. (2013) proposed that impairments in inhibition might be a defining feature of DD. We investigated response inhibition using the stop signal task, given that Szűcs et al. (2013) reported a difference between DD and non-DD children in stop signal accuracy (indeed, this was the only inhibition task where they reported accuracy differences). In addition, we measured sensitivity to interference~~inhibition skills~~ by examining the effect of perceptual congruency on the dot comparison task. Gilmore et al. (2013) and Szűcs et al. (2013b) highlighted the Stroop-like properties of the dot comparison task (i.e., that the congruency between numerosity and the perceptual properties of the stimuli are systematically manipulated), and Gilmore et al. (2013) argued that this property of the task might explain the link between dot comparison performance and mathematics skills. Additionally, Szűcs et al. (2013a) found a greater effect of perceptual congruency in the dot comparison~~is~~ task in a DD group~~than in the case of in a control groups~~ (although see Rouselle & Noël, 2007). Together, the stop signal and dot comparison~~se measures~~~~tasks~~ examined two partially independent aspects of inhibitory control (see Friedman & Miyake, 2004): response control and suppression of interference.

### *Summary*

In summary, this study investigated the core impairments underlying DD. We were particularly interested in whether ordering skills were impaired in DD, and, if this was the case, in the extent and nature of these deficits. In addition to conducting the most comprehensive investigation into order processing skills in DD so far, we also investigated the evidence regarding magnitude processing and inhibition deficits, to get a full picture of the range of difficulties that characterise DD. Within the tasks that we employed, we systematically manipulated some factors (e.g., perceptual congruency, numerical/magnitude

differences, etc.) with the aim of comparing the effects of these manipulations between groups. This way we could investigate the cognitive processes that underlie performance on these tasks in children with and without DD. Finally, we selected the children in the DD group very carefully to make sure that all of these children had persistent and specific difficulties in mathematics, and we used rigorous methods to match the DD and control groups.

## Method

### *Selection of participants*

The children in this study were allocated to a developmental dyscalculia (DD) group or a control group after a two-step screening procedure. First, nineteen schools had been contacted and permission was sought to access the schools' records of children's ( $n = 3,345$ ) performance on age-standardized national curriculum-based mathematics, English and IQ tests. The schools administer standardized tests of mathematics and English every year, starting from year 3. The schools also administer standardized IQ tests, although the timing and frequency of testing varies between schools. Typically, the schools administer IQ tests twice during the primary school years. The standardized achievement tests used by the schools were the *Progress in Maths* (PiM) test, and the *Progress in English* (PiE) test. Additionally, the schools used either (or both) of the following standardized IQ tests: The *Cognitive Abilities Test-Fourth Edition* (CAT4) or the *Non-Reading Intelligence Test* (NRIT).

One hundred-and-twenty children from the original sample, who attended years 5, 6 or 7 of seven different primary schools, were invited to participate in a screening session. The seven schools were selected on the basis that they had a relatively high number of children with a potential diagnosis of DD. Forty children were invited for potential inclusion in the

DD group, and eighty children were selected for potential inclusion in the control group. The children were considered for inclusion in the DD group if their standardized score on the PiM was at least 1 *SD* below the population mean (i.e., they had a score of 85 or lower) for at least two academic years, whereas their PiE score and their IQ score were close to the population mean (i.e., at least 86 for all school years, although children with higher scores were preferred). Children were considered for potential inclusion in the control group if they attended the same schools and classes as the children in the potential DD group, and they were similar in age, gender, and their recent PiE and CAT/NRIT scores. The PiM score of the children in the potential control group was close to the population mean. Children with an official diagnosis of a developmental disorder (other than mathematics difficulties in the case of the DD group) were excluded from the study.

In the second step of the selection procedure, the hundred-and-twenty children who were pre-selected in the first step, were administered an additional standardized measure of mathematics ability, a standardized measure of reading ability, and an IQ test by the researchers (see Materials section). Children from the potential DD group were included in the final DD sample if this additional testing confirmed that they had a standardized score on the maths test that was at least 1 *SD* below the population mean (i.e., a score of 85 or lower), and they had a discrepancy between their IQ and mathematics score of at least 7 standardized points (i.e., a magnitude of .5 *SD* relative to population standards) or a discrepancy of at least 7 points between their standardized mathematics and reading scores, as measured by the researchers<sup>2</sup>. We aimed to recruit children with a relatively large discrepancy between their

<sup>2</sup> The DSM-5 diagnostic criteria for dyscalculia/specific learning disorder in mathematics (in contrast with the DSM-IV criteria), do not require a discrepancy between maths scores and IQ. However, for the purposes of this study, we recruited children with a significant maths-IQ discrepancy, as this can help in disentangling the effects of low maths scores vs. low IQ on their performance on the tasks. We also did this so that we could obtain samples with IQs in the normal range. In terms of the individual profiles of the children in the DD group, in the case of 10 children, there was a discrepancy of at least 7 standard points between both the children's IQ and reading scores and their maths scores (i.e., half of the children only had a difficulty in maths). There were 4 additional children who had a relatively large discrepancy between their IQ and maths scores, but only a

Formatted: Font: Italic

Formatted: Font: Not Italic

Formatted: Font: (Default) +Body (Calibri), 10 pt

mathematics and IQ/reading scores (see Table 1), in order to make sure that children in the DD group had specific difficulties in mathematics. Children from the potential control group were excluded from the final sample if they had standardized mathematics scores under 90 from any academic year (based either on the scores collected by the schools or the researchers). We did this to make sure that none of the children in the control group had any hint of mathematics difficulties.

### *Participants*

On the basis of the selection criteria, twenty children (7 girls) were included in the final DD group. Twenty additional children (10 girls) were included in the control group, who were matched to the children in the DD group as closely as possible on age, gender, IQ and standardized reading scores. Additionally, these children were selected from the same schools and classes as the children with DD. The characteristics of the two groups are presented in Table 1. There was one child in the DD group, and two children in the control group who did not speak English as their first language. Nevertheless, the researchers judged that these children had appropriate English language skills to be able to participate in the study. Regarding the socio-economic status of the children, there were fifteen children in each group who were eligible for free school meals. The most common reason for children to be eligible for free school meals is that their parent/guardian experiences economic hardship. Thus, on

---

small discrepancy between their maths and English scores (i.e., they had difficulties with both maths and English). Nevertheless, these children did not have a diagnosis of dyslexia, and, for this reason, we decided to keep them in our dyscalculia sample. In the case of the remaining 6 children, there was a large discrepancy between their maths and English scores, but only a small discrepancy between their maths and IQ scores. We decided to include these children in our sample on the basis that their sustained difficulties with maths did not extend to other aspects of learning. Although we acknowledge that this results in a DD group with a somewhat heterogeneous cognitive profile, this is quite common in the literature on dyscalculia. Regarding the profile of the control children, we acknowledge that they did not represent well a typical population. However, our aim was to maximize the differences in maths skills, while minimizing the differences in all other relevant factors between the children in the two groups.

In the case of 6 participants, the discrepancy between their IQ and maths scores was less than 7 points. Nevertheless, we decided to retain these children in the DD sample on the basis that they showed a significant discrepancy between their reading and maths scores. Our reasoning was that we wanted to identify children who did not have a general intellectual disability (which would correspond to a separate diagnostic

average, children in the sample had a relatively low socio-economic status. The primary schools that the children were recruited from represented a mix of urban schools and outlying rural schools. The study received approval from the School of Psychology research ethics committee. Informed consent was obtained from the schools and the parents of all children who participated. Additionally, verbal assent was obtained from each child at the start of each testing session.

### *Materials*

*Standardized mathematics test:* the *Mathematics Assessment for Learning and Teaching* test (MaLT; Williams, 2005) is a group-administered written test. The MaLT test was developed in accordance with the National Curriculum and National Numeracy Strategy for England and Wales. Test items cover: counting and understanding number (12 items – these tasks included counting forward or backward in steps of constant size – e.g., 5 9 17 21 25), knowing and using number facts (2 items which included deriving and recalling facts related to arithmetic operations – e.g.,  $(5 \times 7) - = 29$ ), calculating (11 items – e.g., “What number is ten less than 4000?”), understanding shape (6 items – e.g., finding the missing piece of a jigsaw based on a picture), measurement (7 items – e.g., finding the area of polygon presented in a grid), and handling data (7 items – interpreting data presented in graphs or pictograms - e.g., using a daily temperature chart, answering the following question: “How many days was the temperature less than 11°C?”). This test allows for invigilators to read the questions to the children, if required, to ensure that test performance reflects mathematics ability rather than reading proficiency. The MaLT test was standardized in 2005 with children from 120 schools throughout England and Wales (MaLT 9,  $\alpha = .93$ ; MaLT 10,  $\alpha = .92$ ). Children had to complete the test within 45 minutes.

---

Insert Table 1 about here

---

*Standardized reading test: the Hodder Group Reading Test II* (HGRT II; Vincent & Crumpler, 2007) is a group-administered multiple-choice test that assesses children's reading of words, sentences and passages. The test was standardized in 2005 with children from 111 primary and secondary schools throughout England and Wales (HGRT-II, level 2,  $\alpha = .95$ ). The test has two parallel forms, which were used in the current project to minimise copying. Children had to complete the test within 30 minutes.

*IQ test:* a short form of the *Wechsler Intelligence Scale for Children-fourth edition* (WISC-IV UK; Wechsler, 2003) that consisted of the block design (non-verbal) and vocabulary (verbal) subtests was individually administered to each participant. This combination of subtests has the highest validity and reliability of the two-subtest short forms of the WISC, and, on the basis of these tasks, full-scale IQ scores can be estimated using the method outlined by Sattler and Dumont (2004).

#### *Order processing tasks*

A short questionnaire, the *Parental Report of Everyday Ordering Ability*, was administered to obtain parents' ratings of their child's ability to perform everyday tasks that involve order processing. The questionnaire was an adapted version of the scale used by O'Connor et al. (2017). As the original scale was developed for 4-to-5-year-olds, some items were modified to make them appropriate for older children. The parents were asked to rate on a 7-point Likert scale (ranging from 1 = *totally disagree* to 7 = *totally agree*) a list of seven statements referring to everyday ordering tasks, such as remembering schedules, planning sequential actions, and carrying out tasks in the appropriate order. Example questions are: "My son/daughter can easily adjust to changes in routine." or "My son/daughter is able to plan a sequence of activities independently" (see Appendix for the full questionnaire). A sum score was computed for the scale. In the present sample, Cronbach's alpha was .85.

The *order working memory task* (the animal race), based on the task used by Majerus, Poncelet, Greffe, and van der Linden (2006), was administered to measure working memory for order. The stimuli used for this task were seven monosyllabic animal names: cat, dog, sheep, bird, bear, horse and fish. Only names with high lexical frequency were used to minimise the requirement to remember the names. The seven animal names were used to form lists with lengths ranging from two to seven items, and there were four trials at each list length. The items were selected randomly from the pool of seven items, and no item could occur twice in the same list. The stimuli were presented by increasing list length. Each stimulus list was presented via headphones, with the experimenter activating each list presentation. After the auditory presentation of the list of animal names, the child was given cards, depicting the animals in the list. Thus, for example, for list length 2 the child received two cards depicting the animals in the list. The child then had to arrange the cards (given in alphabetical order) to match the order of presentation of the auditory sequence by putting them on a cardboard sheet with a staircase-like platform. The child was told whenever the list length increased. Testing was terminated when more than two lists of the same length were reconstructed incorrectly. A point was given to the child for each list that was correctly reconstructed, with every item placed in the appropriate position. [These scores were summed to obtained a total correct score.](#) Additionally, a span score was computed, based on the length of the list that the child was able to recall. To have a span score of 3, a child had to be able to recall at least two out of the four lists with three items correctly. In the present sample, the split-half reliability of the total score (using the Spearman-Brown formula to compare performance on the first two vs. the second two trials at each set length) was .92.

A computerized version of the backward matrices task was administered to measure visuospatial working memory (based on the task used by Mammarella et al., 2015). The task involves remembering the location and the order of blue squares displayed sequentially in a

4x4 grid on a computer screen. The children were asked to recall the location of the blue squares in *reverse order*. The number of squares in the sequence was successively increased from two to eight. Two practice trials were presented, followed by 14 experimental trials (two trials at each sequence length). Children worked through all trials regardless of their performance. A point was given to the child for each sequence that was correctly recalled, and these scores were summed to obtain a total score. Additionally, a span score was computed, based on the number of locations that the child was able to recall in the correct order. To have a span score of 3, a child had to be able to recall at least one out of the two three-item sequences correctly (if a child failed both trials with 3 items, but was able to complete at least one 4-item trial correctly, they were allocated a span score of 4). Nevertheless, we only used span scores to make it easier to interpret the performance of the children at the group levels. We used total scores for our main analyses, as we considered these more reliable. In the present sample, the split-half reliability (using the Spearman-Brown formula to compare performance on the first vs. second trial at each sequence length) was .68.

#### *Ordinal judgment tasks*

Two computerized tasks were administered. One of these was designed to measure *number ordering ability*<sup>3</sup>. In this task, children were presented with number triads (e.g., 2 3 7), and they were asked to decide whether the triads were in the correct increasing order from left to right, irrespective of the numerical distance between the numbers. All numbers were between 1 and 9. Four practice trials were presented, followed by 48 experimental trials. Cronbach's alpha for our sample was .92 for accuracy and .86 for RT.

An additional task, specifically designed for the current study, was administered to measure *annual event ordering* (based on Friedman, 2002). Children were presented with

---

<sup>3</sup> [This task was modelled on Morsanyi et al. \(2017\). A list of all trials can be found in the paper.](#)

three pictures referring to special days or events during the calendar year (e.g., Easter, Halloween and Christmas), and were asked to indicate whether the triads were in the correct order from left to right, the way these events happen during one calendar year. The nine events included were Valentine's day, Easter, sports day (held by the schools in June every year), summer holiday, going back to school after the summer holiday, Halloween, Christmas, New Year's Eve and the child's birthday. As in the number ordering task, four practice trials were presented, followed by 48 experimental trials. Cronbach's alpha was .89 for accuracy and .91 for RT.

The trials in the number ordering and annual event ordering tasks were matched, so that for each number ordering trial there was a corresponding event ordering trial. The number triad 2 3 8 for example, might correspond to the event triad of Easter, sports day and Christmas because Easter is the second event happening during the year, sports day is the third event and Christmas the eighth event. (Depending on the child's birthday, different versions of the task were created to ensure the correspondence between the number and event ordering tasks.) The trials in both tasks were designed so that the distance between the two extreme numbers/events within each triad was systematically manipulated. The smallest distance was 2 (three consecutive numbers/events) and the largest distance was 7. For each distance, eight trials were presented in each task (four in the correct order and four in mixed order). Triads including the two extreme numbers and events were not included to avoid the possibility that participants respond to these trials using a response rule that does not necessitate the on-line checking of order within the triads. Additionally, to ensure that participants processed all three items within the triads, in the case of mixed order triads the first two items always appeared in the correct order.

*Magnitude comparison and estimation tasks*

A computerized task was administered to measure *symbolic number comparison* ability (based on the task used by Dehaene, Dupoux, & Mehler, 1990). Children were presented with two one-digit numerals between 1 and 9, one on the left side of the screen and the other one on the right side, and they were asked to indicate which number was larger. The distance between the two numbers was systematically manipulated between 1 and 6. For each distance, eight trials were presented (resulting in a total of 48 experimental trials). The task started with four practice trials. Cronbach's alphas for both accuracy and RTs were .96.

Another computerized task was developed to measure *non-symbolic comparison ability* (modelled on the task used by Price et al, 2007). Children were presented with two arrays of blocks, one array on the left side of the screen and one on the right side, and they were asked to indicate which array contained more blocks. Each array contained between 1 and 9 blocks. The stimuli were created in such a way that continuous quantity variables such as area and density of the squares could not be reliably used to select the correct array. Additionally, on half of the trials, the array with more blocks had a bigger total surface area (congruent trials), and in the other half the array with more blocks had a smaller total surface area (incongruent trials). The numerical difference between the two arrays was systematically manipulated between 1 and 6, and, for each difference, eight trials were presented. The task was presented as a game in which the arrays showed how many sweets each of two children had (this was modelled on Gilmore, Attridge, De Smedt, & Inglis, 2014). The arrays were presented for 1 second, after which they disappeared from the computer screen. Once the blocks disappeared from the screen, the children were prompted to indicate which of the two characters had more sweets. Four practice trials were presented, followed by 48 experimental trials. In the present sample, Cronbach's alpha was .50, which is poor, but not unacceptable (Kline, 2000).

A computerized version of the *number line estimation task* was administered to assess children's ability to spatially represent numbers along a mental number line. This task was

based on the number-to-position problems used by Siegler and Opfer (2003). In each problem, children were presented with a number and asked to estimate where it would appear on a number line by using the mouse to click on the line. We used 2 different scales for the task: a 0-100, and a 0-1000 scale. Both scales were 1000 pixels long, which made the errors of estimation (i.e., the distance in pixels between the correct position of the target number and children's estimation of the position of the number) across scales directly comparable. The task included 10 problems for each scale. The numbers presented were 2, 3, 4, 6, 18, 25, 42, 67, 71 and 86 for the 0-100 scale and 4, 6, 18, 25, 71, 86, 230, 390, 780, 810 for the 0-1000 scale. Errors were averaged across all 10 trials for each scale, with a higher number indicating worse performance. Cronbach's alpha for the estimation errors was .79.

#### *Inhibition*

The *stop signal task* (Logan & Cowan, 1984) was administered to measure response inhibition. In the task, a smiley face was presented on a black background in the middle of the screen. The smiley face was followed by a white arrow, which was pointing left or right. The presentation of the arrow was followed by either a sound (the stop signal) or no sound. Children were required to indicate the direction of the arrow using a key press during go trials, and to withhold their response during stop trials. A block of 30 go trials was administered first to calculate the mean RT of the child. This was followed by 140 alternating go and stop trials, which were presented in three blocks: a practice block with 28 trials, and two experimental blocks with 56 trials each. The stop signal was presented at four different intervals: 200, 300, 400, and 500 ms before the mean RT of the child, calculated on the basis of the first block of go trials. In the case of go trials, children had to respond within a 1000 ms time window. Twelve trials of each interval were administered in total in the last two blocks, leading to a ratio of go and stop trials of 2.33:1. Mean accuracy of the stop trials in

blocks 3 and 4 was calculated as a measure of performance. Cronbach's alpha for our sample was .84.

We also administered a *choice reaction time task* (based on Fry & Hale, 1996) to measure basic processing speed, given that some of our other tasks involved RT measures. Children were asked to press a red or a blue button in response to the presence of a red or a blue circle on the computer screen. Forty trials were administered (half red, half blue). Performance was indexed by children's mean RT for correct trials. Cronbach's alpha for the present sample was .94.

#### *Procedure*

The participants completed the tasks in three testing sessions. The order in which the tasks were presented was the same for all participants. The first session was a group session, which took approximately 80 minutes. In this session, the participants were administered the MaLT and the HGRT-II. The other two sessions were individual sessions of about 35 minutes. In the first individual session, the children completed the following tasks (in this order): backward matrices, choice RT, annual events ordering, symbolic comparison, and order working memory. The second individual session started with the number line estimation task. This was followed by number ordering, the stop signal task, non-symbolic comparison, and the two subtests of the WISC-IV: block design and vocabulary. On average, there were 7 days between the group session and the first individual session, and two days between the first individual session and the second individual session.

#### Results

Descriptive statistics for overall performance on each task are displayed in Table 2. As a preliminary analysis, we compared choice reaction times between the DD and control groups (including performance on the correct trials only). Mean choice RTs were not

significantly different between groups ( $p = .130$ ). Nevertheless, as there was a trend towards shorter RTs in the control group, we included choice RT as a covariate in some of our analyses.

#### *Order processing tasks*

First we analysed the results regarding the *parental order processing questionnaire*. The average score of the children in the DD and control groups were significantly different [ $t(38) = 2.42$   $p = .021$ , Cohen's  $d = .76$ ], indicating worse everyday ordering skills in the case of children with DD<sup>4</sup>.

Then we investigated performance on the order memory tasks. In the case of the *order working memory task*, children with DD had an average serial order memory span of 3.85 items. Children in the control group had an average order memory span of 4.40 items. The difference between groups in the total number of correctly recalled sequences was significant,  $t(38) = 2.78$   $p = .008$ , Cohen's  $d = .88$ .

In the case of the *backward matrices task*, children with DD had an average visual-spatial working memory span of 5.75, whereas the control children had an average order memory span of 6.05 items<sup>5</sup>. There was a significant difference between groups in the total number of correctly recalled sequences,  $t(38) = 2.13$   $p = .040$ , Cohen's  $d = .76$ .

In the case of this task, it was also possible to compare the number of correctly recalled item locations between the two groups (i.e., when children recalled the right locations within a sequence, but not in the correct order). Out of 70 item locations, DD participants correctly recalled on average 62.20 ( $SD=2.88$ ), and the control participants recalled 62.45 ( $SD=2.70$ ), which showed that accuracy was high (about 89%) and not significantly different between the

<sup>4</sup> It could be argued that a more appropriate test to analyse these results would be a Mann-Whitney U test. This analysis also yields a significant difference between the groups ( $p=.038$ ).

<sup>5</sup> The span score was higher than

two groups ( $p=.779$ ). That is, there was a group difference in memory for order, but not in item memory on this task.

We then analysed performance on the *order judgment tasks*. We included the results of the number and annual event ordering tasks in the same analyses because we expected similar distance effects on each task. However, we considered the results separately for correct order and mixed-order trials (i.e., trials in which items were not shown in the correct order), because we expected the distance effects to be different across these trials. This grouping of the results was supported by the strong correlations between performance on the two tasks, both for correct order trials [ $r(38) = .61, p < .001$ ] and for mixed-order trials [ $r(38) = .42, p = .007$ ], whereas the correlations between correct order and mixed-order trials within each task were weaker and non-significant at this sample size ( $rs < .19$ ). We analysed the results separately for accuracy and RTs.

Regarding accuracy on the correct order trials (Figure 1), a  $2 \times 6 \times 2$  mixed ANOVA with task (number/event ordering) and distance between the two extreme items in the triad (2/3/4/5/6/7) as within-subjects factors, and group (DD/control) as a between-subjects factor indicated a significant effect of distance [ $F(3.32, 126.13) = 6.26, p < .001, \eta_p^2 = .14$ ], and a significant effect of task [ $F(1, 38) = 14.59, p < .001, \eta_p^2 = .28$ ], but no significant effect of group, and no significant interactions ( $ps > .190$ ). The effect of task was present because accuracy was higher in the case of the number ordering task ( $M = .80, SD = .29$ ) than in the case of the annual event ordering task ( $M = .65, SD = .29$ ). Follow-up analyses using Bonferroni-Holm corrections indicated that accuracy for trials with a distance of 2 was significantly higher than for trials with distances of 4, 6 and 7. All other contrasts were non-significant.

A similar analysis of RTs showed only a significant effect of task [ $F(1,38) = 14.59, p < .001, \eta_p^2 = .28$ ], with shorter RTs for the number ordering task ( $M = 2911$  ms,  $SD = 1185$  ms) than for the annual event ordering task ( $M = 3981$  ms,  $SD = 1902$  ms).

In summary, both the accuracy and RT results indicated that participants found the annual event ordering task more difficult than the number ordering task. Additionally, the accuracy results indicated a reverse distance effect in the case of both tasks.

---

Insert Figure 1 about here

---

Regarding accuracy on the mixed-order trials (Figure 2), a 2x6x2 mixed ANOVA with task (number/event ordering) and distance between the two extreme items (2/3/4/5/6/7) as within-subjects factors, and group (DD/control) as a between-subjects factor indicated a significant effect of distance [ $F(5,190) = 9.58, p < .001, \eta_p^2 = .20$ ], a significant effect of task [ $F(1,38) = 21.18, p < .001, \eta_p^2 = .36$ ], and a significant effect of group [ $F(1,38) = 9.48, p = .004, \eta_p^2 = .20$ ]. There was also a significant distance by task interaction [ $F(5,190) = 3.50, p = .005, \eta_p^2 = .08$ ]. None of the other interaction effects were significant ( $ps > .540$ ). The effect of task was present, because accuracy was higher in the case of the number ordering task ( $M = .93, SD = .08$ ) than in the case of annual event ordering ( $M = .84, SD = .14$ ). The group effect indicated that accuracy was lower in the DD group ( $M = .85, SD = .10$ ) than in the control group ( $M = .93, SD = .06$ ).

In order to follow up on the distance by task interaction, we ran separate one-way ANOVAs for each task with distance (2/3/4/5/6/7) as a within-subjects factor. In the case of the number ordering task, there was a significant effect of distance [ $F(3.67, 143.28) = 2.70, p = .037, \eta_p^2 = .07$ ]. Follow-up analyses with Bonferroni-Holm corrections indicated that

accuracy for trials with a distance of 2 was significantly lower than the accuracy of trials with a distance of 6.

Regarding the annual event ordering task, a similar analysis indicated a significant effect of distance ( $F(5, 195) = 8.28, p < .031, \eta_p^2 = .18$ ). Follow-up analyses with Bonferroni-Holm corrections indicated that accuracy for trials with a distance of 2 were significantly lower than for trials with a distance of 6. Additionally, accuracy for trials with a difference of 3 was significantly lower than for trials with distances of 4, 5, 6 and 7.

---

Insert Figure 2 about here

---

In summary, the accuracy results on the mixed-order trials showed that participants found the annual event ordering task more difficult than the number ordering task, which was similar to the results regarding correct order trials. There was also a significant *canonical* distance effect on both tasks, with a stronger effect in the case of annual event ordering. Additionally, DD children performed more poorly on both tasks.

We also analysed RTs on the mixed-order trials. A 2x6x2 mixed ANOVA with task (number/event ordering) and distance (2/3/4/5/6/7) as within-subjects factors, and group (DD/control) as a between-subjects factor indicated a significant effect of task ( $F(1,38) = 15.56, p < .001, \eta_p^2 = .29$ ), with shorter RTs for number ordering ( $M = 2847$  ms,  $SD = 1112$  ms) than for annual event ordering ( $M = 3822$  ms,  $SD = 1809$  ms). There was also a significant task by group interaction [ $F(1,38) = 5.34, p = .026, \eta_p^2 = .12$ ], and a marginal effect of distance ( $p = .083$ ). The task by group interaction remained significant ( $p = .022$ ) when the analysis was re-run with choice RT as a covariate. Follow-up analyses with Bonferroni-Holm corrections indicated that the RT difference between groups was significant in the case of the number ordering task ( $M = 3242$  ms,  $SD = 1179$  ms, for the DD group and

$M = 2452$  ms,  $SD = 904$  ms, in the case of the control group), but it was not significant in the case of the annual event ordering task ( $M = 3635$  ms,  $SD = 1939$ , for the DD group and  $M = 4008$  ms,  $SD = 1699$  ms, in the case of the control group). When choice RT was included as a covariate in the analysis of group differences on the number ordering task, the effect was slightly reduced ( $p = .044$ ), and was no longer significant after applying Bonferroni-Holm corrections. The RT results on the mixed-order trials were in line with the previous results in showing that participants found the annual event ordering task more difficult than the number ordering task. Additionally, the task by group interaction indicated that children in the DD group tended to respond slower in the case of number ordering trials than controls, whereas the two groups had similar RTs in the case of the annual event ordering task.

#### *Magnitude comparison and estimation tasks*

We analysed accuracy on the *symbolic comparison task* using a 6x2 mixed ANOVA with numerical distance (1/2/3/4/5/6) as a within-subjects factor, and group (DD/control) as a between-subjects factor (Figure 3). There was a significant main effect of distance [ $F(1.49,56.63) = 6.92$   $p < .001$   $\eta_p^2 = .15$ ], but no effect of group ( $p = .253$ ), and no distance by group interaction ( $p = .538$ ). Follow-up analyses, using Bonferroni-Holm corrections indicated that there was a significant difference between accuracy for items with a numerical distance of 1 and items with a numerical distance of 4. Additionally, there was a significant difference between items with a numerical distance of 2 and items with numerical distances of 4, 5 and 6.

\_\_\_\_\_  
Insert Figure 3 about here  
\_\_\_\_\_

We also analysed RTs on the *symbolic comparison task* using a 6x2 mixed ANOVA with numerical distance (1/2/3/4/5/6) as a within-subjects factor, and group (DD/control) as a between-subjects factor. There was a significant main effect of group [ $F(1, 38) = 5.16$   $p = .029$   $\eta_p^2 = .12$ ], with children with DD responding more slowly ( $M = 1276$  ms,  $SD = 552$  ms) than controls ( $M = 964$  ms,  $SD = 267$  ms). There was also a marginal effect of distance ( $p = .083$ ), but there was no distance by group interaction ( $p = .757$ ). When we repeated the ANOVA analysis including choice RT as a covariate, we found that the effect of choice RT was significant [ $F(1, 37) = 5.68$   $p = .022$   $\eta_p^2 = .13$ ], whereas the effect of group was reduced to marginal ( $p = .088$ ).

We analysed accuracy on the *nonsymbolic comparison task*, using a 6x2x2 mixed ANOVA with numerical distance between the displays (1/2/3/4/5/6) and congruency between the surface area of the objects and numerosity (congruent/incongruent) as within-subjects factors, and group (DD/control) as a between-subjects factor (Figure 4). There was a significant effect of distance [ $F(3.23, 122.69) = 28.59$ ,  $p < .001$ ,  $\eta_p^2 = .43$ ] and a significant effect of group [ $F(1, 38) = 8.10$ ,  $p = .007$ ,  $\eta_p^2 = .18$ ], with lower accuracy in the DD group ( $M = .91$ ,  $SD = .06$ ) than in the control group ( $M = .95$ ,  $SD = .03$ )<sup>6</sup>. There was also a marginal effect of congruency ( $p = .051$ ), but the interaction effects were not significant ( $ps > .20$ ). Pairwise comparisons using Bonferroni-Holm corrections showed that accuracy for trials with a numerical distance of 1 was significantly lower than for trials with numerical distances of 2, 3, 4, 5 and 6. Additionally, accuracy for trials with a numerical distance of 2 was significantly lower than for trials with numerical distances of 4 and 5. All other comparisons yielded nonsignificant results.

---

<sup>6</sup> It could be argued that, because performance was close to ceiling on this task, the ANOVA analyses were not appropriate. Nevertheless, the group difference was also present when we used a Mann-Whitney U test ( $p = .005$ ). We preferred to present the results of the ANOVA analyses in the main text, because this is the typical analysis strategy in the relevant literature (e.g., Gilmore et al., 2014; Price et al., 2007).

Formatted: Font: Italic

Formatted: Font: (Default) +Body (Calibri), 10 pt

We also analysed performance on the *number line task*. We ran a 2x2 mixed ANOVA to investigate the effect of scale type (0-100/0-1000) as a within-subjects factor and group (DD/control) as a between-subjects factor. There was a significant effect of scale type [ $F(1, 38) = 70.38$   $p < .001$   $\eta_p^2 = .65$ ], which was present, because the average estimation error for the 0-100 scale ( $M = 59.65$ ,  $SD = 36.46$ ) was smaller than for the 0-1000 scale ( $M = 148.24$ ,  $SD = 73.09$ ). Additionally, there was also a significant effect of group [ $F(1, 38) = 7.79$   $p = .008$   $\eta_p^2 = .17$ ] with larger average error in the DD group ( $M = 123.19$ ,  $SD = 48.55$ ) than in the control group ( $M = 84.70$ ,  $SD = 38.02$ ). The scale type by group interaction was non-significant ( $p = .452$ ).

---

Insert Figure 4 about here

---

### *Inhibition*

In the case of the stop signal task, we first compared the DD and control groups on go trial accuracy in the case of blocks 3 and 4, where stop and go trials were mixed to check that children in both groups engaged with the task (i.e., that they pressed the button within the allocated time window of 1000 ms). Mean accuracy in the DD group was .80 ( $SD = .14$ ) and it was .83 ( $SD = .22$ ) in the control group ( $p = .622$ ). We analysed the results regarding the “stop” trials using a mixed 2x4 ANOVA with stop signal delay (200/300/400/500 ms) as a within-subjects factor, and group (DD/control) as a between-subjects factor (see results in Figure 5). There was a significant effect of tone delay [ $F(2.37, 89.88) = 17.24$   $p < .001$   $\eta_p^2 = .31$ ], but no main effect of group, and no interaction between tone delay and group ( $ps > .35$ ). Follow-up analyses showed that performance at 200 ms delay was significantly worse than performance at all other delays, and performance at 300 ms delay was significantly worse

than performance at 500 ms delay. The other comparisons were non-significant after applying a Bonferroni-Holm correction.

---

Insert Figure 5 about here

---

#### *Best predictors of group membership*

For diagnostic purposes, it is important to identify tasks that can reliably discriminate between individuals with and without DD. Given that there were several tasks that reliably discriminated between the two groups, we were interested in selecting the best predictors of group membership from this set of tasks. For this reason, we carried out a stepwise logistic regression analysis (using the forward conditional method), with group membership (DD/control) as the dependent variable, and [total scores on the parental order processing questionnaire, the order working memory and backward matrices tasks, as well as a combined-accuracy scores on the mixed-order trials on the daily events, and number ordering and dot comparison tasks, the animal race task, the backward matrices task, the dot comparison task, and the and average error on the number line task](#) as predictor variables (i.e., we only considered tasks where we found a significant group difference). The forward conditional method enters variables on the basis of their level of significance (Wald), and performs removal testing based on the probability of a likelihood-ratio statistic based on conditional parameter estimates (i.e., variables are removed if this does not result in a significant drop in the proportion of variance explained by the model). [The stepwise method was chosen because we wanted to obtain a model that only included significant predictors. This method also offers a way to deal with multicollinearity as the contribution of each](#)

variable is considered one by one (Tabachnick & Fidell, 2007). In our model, multicollinearity was a potential problem, because of the strong correlation between the number and event ordering tasks<sup>7</sup>. Table 42 presents the order of entering the predictors into the regression equation, the contribution of each variable characteristics of each model, and change in the explained variance if a variable is removed, comparisons between the models. Comparisons between the models confirmed that adding each variable resulted in a significant increase in the variance explained by the models. The final model, which explained the greatest proportion of variance (Nagelkerke  $R^2 = .63$ ), included the following predictors: order working memory, parental questionnaire and the number line task. This model categorised 80% of the participants correctly as DD/control (Table 5). Interestingly, model 4, which only included order working memory and the parental questionnaire, performed even better in categorising the participants (with 82.5% categorised correctly), although the amount of variance explained by this model (Nagelkerke  $R^2 = .48$ ) was significantly lower than the variance explained by model 5<sup>8</sup>.

Formatted: Font: 12 pt

---

Insert Tables 42 and 5 about here

---

---

Insert Table 3 about here

---

---

---

<sup>7</sup> Another way to deal with multicollinearity would be to combine the scores from highly correlated variables. However, for practical purposes, it is more useful to assess the individual contribution of each task.

<sup>8</sup> The results regarding explained variance and the model's ability to identify participants as DD/control might seem contradictory. It might be useful to consider that model 4 was particularly successful at identifying participants who were *not* dyscalculic, whereas model 5 showed the best performance in identifying DD participants (i.e., correct classification is not additive, when a new variable is included in the model). In other words, overall explained variance in diagnostic status increases with each additional relevant predictor variable, however, the ability of the model to correctly identify an individual's diagnostic status will depend on the fit between each individual's profile and the profile predicted by the statistical model.

Formatted: Font: Italic

## Discussion

### *Results regarding order processing*

The main aim of this study was to test the hypothesis that DD involves order processing deficits using the broadest range of order processing measures so far. These included some novel measures that have never been used in the case of participants with DD before: a questionnaire on everyday ordering abilities, and an annual event ordering task. Our results supported the hypothesis that children with DD have order processing difficulties, including problems with everyday activities that require ordering (e.g., recalling the order in which past events happened), recalling short, novel sequences (of both verbal materials and spatial locations), and making judgments about the correctness of the order of familiar sequences (including both numerical and non-numerical items) that draw on long-term memory representations.

In the case of memory for unfamiliar/novel sequences, the verbal task did not require memory for items (i.e., the task was a pure measure of order memory). The spatial task required memory for locations and order as well. In the case of this task, we have confirmed that children with DD recalled locations just as accurately as controls, but they showed impaired memory for order. This extends the findings of Attout and Majerus (2014) who showed impaired order memory, but intact item memory for verbal materials in DD.

With regard to order judgments about familiar sequences, our materials were designed in such a way that a direct comparison was possible between numerical and non-numerical sequences. Earlier studies showed similar (reverse) distance effects in the case of ordered sequences of numbers, letters and months (e.g., Jou & Aldridge, 1999; Goffin & Ansari, 2016; Lyons & Beilock, 2013; Morsanyi et al., 2017; Turconi, Campbell, & Seron, 2006; Vos et al., 2017). When present, canonical distance effects had been found in the case of mixed-order sequences of both sorts of familiar items, and reverse distance effects had been found in

the case of correctly ordered sequences<sup>9</sup>. In order to investigate these effects, we analysed distance effects on correctly ordered and mixed-order trials separately. A key question regarding these tasks was whether there was any evidence that DD children show specific deficits in numerical order judgments (which are distinct from their order judgments regarding annual events).

As expected, children with DD performed more poorly on these order judgment tasks than controls. Specifically, they were less able to recognize when the order of items within a triad was incorrect. Importantly, this effect was present in the case of both number and event ordering. Additionally, although the annual events task was more difficult for children, as indicated by both the accuracy and RT results, the distance effects were remarkably similar across tasks, with reverse distance effects in the case of correct order trials, and canonical distance effects in the case of mixed-order trials. Importantly, the distance effects were also similar across groups, indicating that the underlying cognitive processes in the two groups were not qualitatively different. The only result that hinted at any domain-specific effect was the finding that the RT difference between groups was larger in the case of mixed-order trials on the number ordering task than in the case of the annual event ordering task. Nevertheless, thise group difference was not reliable after controlling for basic RT.

In summary, the results of the ordering tasks showed strong evidence for ordering deficits in DD (indeed, this effect was absent only in the case of correctly ordered trials in the number and event ordering tasks). The results were similar across a broad range of measures. Importantly, most of these did not include numbers, and when number ordering was compared to an annual event ordering task, the results were very similar, both in terms of

---

<sup>9</sup> These effects indicate that in the case of correctly ordered trials, participants find it easier to recognize correct trials when the items immediately follow each other (e.g., 1 2 3), but they sometimes incorrectly reject trials where the items do not form a familiar sequence (e.g., 2 5 8). In the case of mixed-order trials, participants experience more difficulty on close trials (e.g., 2 4 3) – i.e., sometimes they incorrectly accept these.

Formatted: Font: (Default) +Body (Calibri), 10 pt

group differences across the two tasks and distance effects. Thus, these results strongly suggest that ordering deficits in DD are not restricted to the domain of numbers ([see also Attout, & Majerus, 2014; Attout et al., 2015; De Visscher et al., 2015; Rubinsten & Sury, 2011](#)).

#### *Magnitude comparison and estimation*

In addition to the investigation of ordering ability in DD, we also administered a series of tasks to measure magnitude comparison and estimation skills. There were clear group differences in the case of two tasks: the non-symbolic comparison task and the number line task. The non-symbolic comparison task that we used was similar to the task used by Price et al. (2007), although our task was not designed to investigate RTs. We found similar distance effects on the task as Price et al. (2007). However, we did not find a group by distance interaction. Instead, we found a main effect of group, with DD children performing more poorly on the task across all distances. As the distance effect was large, it is unlikely that we did not detect the interaction effect due to a lack of statistical power. In fact, the distance by group interaction was very far from being significant ( $p = .651$ ). Thus, as in the case of the ordering tasks, we can assume that although children with DD were less accurate in their magnitude comparison judgments, the underlying cognitive processes were similar across groups.

As we noted in the introduction, there is inconsistency in the literature regarding dot comparison performance in dyscalculia. A possible issue to consider is that there is evidence that the way the stimuli are designed have a strong effect on performance on the task (e.g., Szűcs et al., 2013). Although most of this literature is focussed on the way visual stimulus parameters are manipulated in the task, the numerosity of the dots that are presented might also be important. There is evidence that the estimation of small and larger numerosities is based on different processes. Burr, Anobile and Arrighi (2017) presented

Formatted: Indent: First line: 1 cm, Line spacing: Double

psychophysiological evidence that there are three different mechanisms involved in numerosity processing, depending on the number and perceptual properties of the items. Separate mechanisms are used in the subitizing range (up to about 4 items), in the small numerosity range (from about 5 items up to about 30 items, depending on the density of the display), and when a larger number of items are presented, in particular when these items merge into a texture. According to Burr et al. (2017), only estimation skills in the small numerosity range are related to mathematics abilities in typical participants. Interestingly, Ashkenazi, Mark-Zigdon and Henik (2013) found an impairment in DD in both the subitizing (1-4) and the small number estimation (5-9) ranges, which exactly correspond to the number of items that were used in our study. Unfortunately, these authors did not investigate estimation skills including larger numbers. Nevertheless, based on this evidence, a possible reason for discrepancies in the literature regarding the links between non-symbolic comparison and maths skills might be that the non-symbolic comparison tasks show a great variability in the number of items that they present, and potentially tap into different underlying processes.

The logistic regression results also offer interesting insight into whether this task is useful in discriminating between dyscalculic and control participants. This task was entered first into the regression equation (model 1), on the basis that among all individual predictors, ~~this task had~~ the regression coefficient ~~of associated with~~ this task was associated with the lowest  $p$  value. Nevertheless, looking at the results ~~in~~ Table 5, it can be seen that on the basis of their performance on this task, participants with DD were categorised as dyscalculic with 50% probability (i.e., the model performed at chance)<sup>10</sup>. Thus, the task was somewhat

---

<sup>10</sup> For a comparison, on the basis of the parental questionnaire alone, 70% of the participants could be identified correctly as DD or control. Using the order working memory task or the number line task alone, 67.5% of the participants could be correctly classified.

Formatted: Font: Italic

useful in identifying participants who were ~~not~~NOT dyscalculic, but the results indicate that low performance on this task was only present in some participants with DD (i.e., low performance was not generally characteristic of this group). This is in contrast with the performance of model 4, which included two ordering tasks, and performed very successfully at correctly identifying both DD and control participants.

Formatted: Font: Italic

In the case of the number line task, there was once more a group effect ([see also Skagerlund & Traff, 2016](#)), as well as an effect of scale type. Specifically, children's judgments of the position of numbers were more accurate in the case of the 0-100 scale than in the case of the 0-1000 scale, although the two scales were of equal physical length. This suggests that the difference between the accuracy of judgments across scales was related to children's cognitive representations of the number sequences (i.e., more precise representations in the case of smaller numbers), rather than to the way they had to respond to the task. Importantly, we did not find an interaction effect between scale type and group (the effect was non-significant with a very small effect size), which again indicated that the underlying cognitive processes across groups were similar, although children with DD performed more poorly on these tasks.

A final task related to magnitude processing was the symbolic comparison task. ~~We only analysed the RT results in the case of this task, as accuracy was at ceiling in both groups.~~ Once more, we found similar distance effects in the two groups (see also Soltész et al., 2007), suggesting that they processed the stimuli in the same way. Additionally, there was a difference in RTs between the two groups. Nevertheless, once we controlled for choice RTs, the group difference in RTs was reduced to a non-significant trend. The results regarding the symbolic comparison task are in line with Szűcs et al. (2014), but they are in contrast with Rousselle and Noël (2007) who reported group differences on the symbolic comparison task, both for accuracy and RTs.

In summary, group differences were found in the case of the non-symbolic comparison task and the number line task, whereas the evidence for a difference in symbolic comparison performance was weak. As in the case of the order processing tasks, the group differences appeared to be related to the efficiency of processing, rather than to qualitative differences in the way children with DD and controls responded to the tasks.

### *Inhibition*

Inhibition skills are important for mathematics (e.g., Bull & Scerif, 2001; St Clair-Thomson & Gathercole, 2006; Blair & Razza, 2007), and Szűcs et al. (2014) proposed that impairments in inhibition skills might be a defining feature of DD. Additionally, recent studies (e.g., Fuhs & McNeil, 2013; Gilmore, Attridge, Clayton, Cragg, Johnson et al., 2013; Szűcs, Nobes, Devine, Gabriel & Gebuis, 2013) suggested that the link between performance on non-symbolic comparison tasks and mathematics skills might be attributable to the inhibition component of the task, and could be restricted to incongruent trials. Szűcs et al. (2014) also reported a group by congruency interaction when they compared children with DD and controls, with larger congruency effects in the case of children with DD. For these reasons, we were interested in potential group differences ~~manipulated the~~ in congruency ~~of the trial effects~~ in the non-symbolic comparison task. In contrast with Szűcs et al. (2014), but in line with Rousselle and Noël (2007), we found no significant interaction between congruency and group ( $p = .207$ ), and this effect was not only non-significant, but also small in size ( $\eta_p^2 = .048$ ). Nevertheless, it is important to note that the studies that found that incongruent trials were more predictive of mathematics ability than congruent trials all used non-symbolic comparison tasks with a relatively large number of items within the arrays that they presented. Thus, it is possible that we did not replicate these findings because we used arrays with a small number of items, which might rely on different cognitive processes (Burr et al., 2018).

We also did not find evidence of group differences in response inhibition, as measured by the stop signal task. Indeed, the pattern of responses, and the effect of stop signal delay was very similar across groups. These results are in contrast with Szűcs et al. (2014) who found a difference between groups in stop signal accuracy, as well as a greater effect of congruency in the DD group than in the case of controls.

#### *Ordering vs. magnitude processing*

Rubinsten and Sury (2011) proposed that magnitude and order processing are two independent core systems, and ordering in particular is impaired in dyscalculia. Although in our sample of children with dyscalculia, we have found evidence for both ordering and magnitude processing/estimation deficits, the results of the regression analysis showed that children with dyscalculia could be identified with high precision on the basis of their performance on some ordering tasks alone (specifically, the ~~number ordering/daily events tasks~~ parental ordering questionnaire, and the animal race task). In fact, among the magnitude/estimation tasks, the number line task was the best predictor of group membership. Interestingly, the number line task has been described as a task that combines the need for estimation and magnitude-processing skills with an understanding of ordinality (cf., von Aster & Shalev, 2007; Kaufmann & von Aster, 2012).

Regarding the notion of the independence of magnitude processing and ordering skills, some recent papers showed a close link between these skills (e.g., Lyons & Beilock, 2011; Morsanyi et al., 2017; Sasanguie et al., 2017), although these studies also showed that there was only a partial overlap between ordering and magnitude processing skills. Our finding that ordering and magnitude processing deficits co-occur in DD also hint at the possibility that these skills are inherently related. Nevertheless, order processing problems were the best predictors of a diagnosis of DD.

#### *Domain-specific vs. domain-general deficits in developmental dyscalculia*

Whereas much research has focussed on domain-specific impairments in dyscalculia, recent findings, as well as clinical observations, point to more general deficits that extend beyond the domain of numbers. Specifically, DD has been linked to the impairments of order memory (Attout & Majerus, 2014;<sup>11</sup> Attout et al., 2015) as well as visual-spatial memory (Mammarella et al., 2015; Szűcs et al., 2014). Szűcs et al. (2014) also reported inhibition impairments in non-numerical tasks in DD. Morsanyi et al. (2013) presented evidence for impaired verbal reasoning skills in DD, using transitive reasoning tasks<sup>11</sup>. It is remarkable that, with the exception of the tasks measuring inhibition skills, all of these tasks include the requirement to process information about order. Clinical descriptions of DD also highlight problems with order-processing/sequencing outside the numerical domain.

The current results extend these findings by showing, in particular, that children with DD have ordering difficulties in the domain of time (e.g., remembering the order of past events, carrying out a sequence of actions, making judgments about the order of annual events, etc.). This has important implications for interventions for DD. First, children with DD might not only need support with mathematics, but also with being better orientated in time. Indeed, these skills could be very important for the everyday functioning of individuals with DD. Second, non-numerical order training might be useful in making intervention approaches more diverse, and less stressful to children who already accumulated negative experiences with mathematics. In particular, practicing non-numerical ordering skills might be less intimidating, as children do not associate them with academic success.

#### *Limitations and future directions*

---

<sup>11</sup> An example for a transitive reasoning task is: "If bicycles are faster than aeroplanes, and cars are faster than bicycles, then are cars faster than aeroplanes?" This task involves the requirement to order items along a single continuum (in this case, according to how fast they are), and make judgments about the relative position of the items. Some of the tasks also require participants to accept premises that are unbelievable, which might require the inhibition of beliefs.

The current results provide strong evidence for ordering deficits in DD. However, some limitations of our study should be mentioned. One issue is the size of the groups. Although we made every effort to recruit a large sample of participants, and initially screened over 3,000 children, we were only able to identify 20 children with a profile of very low mathematics skills in the absence of low IQ and significant difficulties with reading. Indeed, these children also had relatively low IQ and reading skills, although their scores were within the normal range.

Another limitation is that, as our main focus was on order processing, it can be argued that we did not pay equal attention to other skills, such as magnitude processing/estimation abilities and inhibition skills. Regarding magnitude processing/estimation, we have specifically selected three tasks that were previously used with DD participants, and that we had good reasons to expect to show performance differences between DD and control children. Indeed, this was confirmed for both the dot comparison and the number line task. Nevertheless, future studies could compare performance on multiple use more challenging versions of the dot comparison task, also including larger varying the number s of items and using different ways to control the perceptual properties of the tasks, in the displays to investigate further investigate both estimation abilities and susceptibility to perceptual distractors in DD. Inhibition skills in DD could also be investigated using Stroop tasks.

Regarding the ordering tasks, the correlations between these tasks ranged from weak and non-significant to very strong. Thus, it is likely that although all of these tasks share an ordering/sequencing component, the underlying processes are only partially overlapping. In fact, within the number and event ordering tasks, there was also a dissociation between performance on correctly ordered and mixed-order trials, and the group difference was only present for the latter. Thus, an important future direction could be to investigate the cognitive processes that are implicated in these tasks further.

It is also important to note that, when selecting the participants (in both the DD and the control group), we have excluded all children with an official diagnosis of a developmental disorder. This was useful for specifically investigating the effects of mathematics difficulties. Nevertheless, dyscalculia is a heterogeneous condition (e.g., Kaufmann et al., 2013; Rubinsten & Henik, 2009) and comorbidity with other developmental disorders is very common. Thus, it is likely that the cognitive profile of some children with DD will differ from the typical profile identified in this study.

### *Conclusions*

This study compared the performance of children with developmental dyscalculia and children without mathematics difficulties on a range of tasks assessing ordering skills, magnitude processing/estimation skills, and inhibition. The two groups were closely matched on age, gender, socio-economic status, educational experiences, IQ and reading ability. The findings revealed differences between the groups both in ordering and magnitude processing abilities. Specifically, both numerical and non-numerical ordering skills were impaired in DD, as well as performance on the dot comparison and number line tasks. Nevertheless, these differences appeared to be quantitative, rather than qualitative, as distance effects, as well as other within-task manipulations had the same effect on the performance of both groups. A logistic regression analysis indicated that a combination of the parental ordering questionnaire, order judgment tasks, order working memory and the number line task could be used to correctly identify 80.5% of the participants as dyscalculic or non-dyscalculic. Indeed, the ordering tasks alone identified the same number slightly larger proportion (82.5%) of participants correctly. This has great significance for the development of novel diagnostic methods for DD. In particular, because even very young children can perform some non-numerical ordering tasks, an early diagnosis of susceptibility to maths difficulties might be possible. The finding that both numerical and non-numerical ordering skills are

impaired in dyscalculia extends the findings of previous investigations (Attout, & Majerus, 2014, Attout et al., 2015; [De Visscher et al., 2015](#); Kauffmann et al., 2011; Rubinsten & Sury, 2011). [Indeed, the current study provided stronger evidence for non-numerical than numerical ordering deficits.](#) We also presented a new parental questionnaire, and a novel task to measure non-numerical ordering skills (i.e., the annual events ordering task) that could be used in future studies. The current findings also open up new avenues for designing interventions for individuals with maths difficulties.

## References

[Ashkenazi, S., Mark-Zigdon, N., & Henik, A. \(2013\). Do subitizing deficits in developmental dyscalculia involve pattern recognition weakness? \*Developmental Science\*, 16, 35-46.](#)

Formatted: Font: (Default) Times New Roman, 12 pt

- Ashkenazi, S., Rubinsten, O., & Henik, A. (2009). Attention, automaticity, and developmental dyscalculia. *Neuropsychology*, 23, 535-540. DOI: 10.1037/a0015347
- Attout, L. & Majerus, S. (2014). Working memory deficits in developmental dyscalculia: The importance of serial order. *Child Neuropsychology*, 21, 432-450. DOI: 10.1080/09297049.2014.922170
- Attout, L., Noël, M. P. & Majerus, S. (2015). The relationship between working memory for serial order and numerical development: A longitudinal study. *Developmental Psychology*, 50, 1667-1679. DOI: 10.1037/a0036496
- 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, 230-247. DOI: 10.1080/87565641.2015.1036993.
- Blair, C., & Razza, R. P. (2007). Relating effortful control, executive function, and false belief understanding to emerging math and literacy ability in kindergarten. *Child Development*, 78, 647-663. DOI: 10.1111/j.1467-8624.2007.01019.x
- Booth, J. L., & Siegler, R. S. (2006). Developmental and individual differences in pure numerical estimation. *Developmental Psychology*, 42, 189-201. DOI: 10.1037/0012-1649.41.6.189
- Booth, J. L., & Siegler, R. S. (2008). Numerical magnitude representations influence arithmetic learning. *Child Development*, 79, 1016-1031. DOI: 10.1111/j.1467-8624.2008.01173.x

Bull, R., & Scerif, G. (2001). Executive functioning as a predictor of children's mathematics ability: Inhibition, switching, and working memory. *Developmental Neuropsychology*, 19, 273-293. DOI: 10.1207/S15326942DN1903\_3

Burr, D. C., Anobile, G., & Arrighi, R. (2018). Psychophysical evidence for the number sense. *Philosophical Transactions of the Royal Society, B*, 373; 20170045.

Formatted: Font: (Default) Times New Roman, 12 pt

Formatted: Font: (Default) Times New Roman, 12 pt

Formatted: Font: (Default) Times New Roman, 12 pt

Butterworth, B. (2005). Developmental dyscalculia. In J. I. D. Campbell (Ed.). *Handbook of mathematical cognition* (pp. 455-467). New York: Psychology Press.

Dehaene, S., Dupoux, E. & Mehler, J. (1990). Is numerical comparison digital? Analogical and symbolic effects in two-digit number comparison. *Journal of Experimental Psychology: Human Perception and Performance*, 16, 626–641. DOI:0096-1523/90/\$00.75

De Smedt, B., & Gilmore, C. K. (2011). Defective number module or impaired access? Numerical magnitude processing in first graders with mathematical difficulties. *Journal of Experimental Child Psychology*, 108, 278-292. DOI: 10.1016/j.jecp.2010.09.003

De Smedt, B., Verschaffel, L., & Ghesquière, P. (2009). The predictive value of numerical magnitude comparison for individual differences in mathematics achievement. *Journal of Experimental Child Psychology*, 103, 469-479. DOI: 10.1016/j.jecp.2009.01.010

Espy, K. A., McDiarmid, M. M., Cwik, M. F., Stalets, M. M., Hamby, A., & Senn, T. E. (2004). The contribution of executive functions to emergent mathematic skills in preschool children. *Developmental Neuropsychology*, 26, 465-486. DOI: 10.1207/s15326942dn2601\_6

- Friedman, W. J. (1990). Children's representations of the pattern of daily activities. *Child Development, 61*, 1399-1412. DOI: 10.1111/j.1467-8624.1990.tb02870.x
- Friedman, W. J. (2000). The development of children's knowledge of the times of future events. *Child Development, 71*, 913-932. DOI: 10.1111/1467-8624.00199
- Friedman, W. J. (2002). Children's knowledge of the future distances of daily activities and annual events. *Journal of Cognition and Development, 3*, 333-356. DOI: 10.1207/S15327647JCD0303\_4
- Friedman, N. P., & Miyake, A. (2004). The relations among inhibition and interference control functions: a latent-variable analysis. *Journal of Experimental Psychology: General, 133*, 101-135. DOI: 10.1037/0096-3445.133.1.101
- Fry, A. F., & Hale, S. (1996). Processing speed, working memory, and fluid intelligence: Evidence for a developmental cascade. *Psychological Science, 7*, 237-241. DOI: 10.1111/j.1467-9280.1996.tb00366.x
- Fuhs, M.W., & McNeil, N.M. (2013). ANS acuity and mathematics ability in preschoolers from low-income homes: Contributions of inhibitory control. *Developmental Science, 16*, 136-148. DOI: 10.1111/desc.12013
- Geary, D. C. (2004). Mathematics and learning disabilities. *Journal of Learning Disabilities, 37*, 4-15. DOI: 10.1177/00222194040370010201
- Geary, D. C. (2011). Cognitive predictors of achievement growth in mathematics: a 5-year longitudinal study. *Developmental Psychology, 47*, 1539-1552. DOI: 10.1037/a0025510
- Gilmore, C., Attridge, N., De Smedt, B., & Inglis, M. (2014). Measuring the approximate number system in children: exploring the relationships among different tasks. *Learning and Individual Differences, 29*, 50-58. DOI: 10.1016/j.lindif.2013.10.004

- Gilmore, C., Attridge, N., Clayton, S., Cragg, L., Johnson, S., Marlow, N. Simms, V. & Inglis, M. (2013). Individual differences in inhibitory control, not non-verbal number acuity, correlate with mathematics achievement. *PLOS ONE*, 8(6), e67374. DOI: 10.1371/journal.pone.0067374
- 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. DOI: 10.1016/j.cognition.2016.01.018
- Hannula, M. M., Lepola, J., & Lehtinen, E. (2010). Spontaneous focusing on numerosity as a domain-specific predictor of arithmetical skills. *Journal of Experimental Child Psychology*, 107, 394-406. DOI: 10.1016/j.jecp.2010.06.004
- Hitch, G. J., & McAuley, E. (1991). Working memory in children with specific arithmetical learning difficulties. *British Journal of Psychology*, 82, 375-386. DOI: 10.1111/j.2044-8295.1991.tb02406.x
- Iuculano, T., Tang, J., Hall, C. W. & Butterworth, B. (2008). Core information processing deficits in developmental dyscalculia and low numeracy. *Developmental Science*, 11, 669-680. DOI: 10.1111/j.1467-7687.2008.00716.x
- Jou, J., & Aldridge, J. W. (1999). Memory representation of alphabetic position and interval information. *Journal of Experimental psychology: Learning, Memory, and Cognition*, 25, 680-701. DOI: 10.1037/0278-7393.25.3.680
- [Kaufmann, L., Mazzocco, M. M., Dowker, A., von Aster, M., Goebel, S., Grabner, R., ... & Rubinsten, O. \(2013\). Dyscalculia from a developmental and differential perspective. \*Frontiers in Psychology\*, 4, 516. DOI: 10.3389/fpsyg.2013.00516](#)
- Kaufmann, L., Vogel, S. E., Starke, M., Kremser, C., & Schocke, M. (2009). Numerical and non-numerical ordinality processing in children with and without developmental

Formatted: Font: (Default) Times New Roman, 12 pt

Formatted: Font: (Default) Times New Roman, 12 pt

Formatted: Font: (Default) Times New Roman, 12 pt, Font color: Auto

dyscalculia: Evidence from fMRI. *Cognitive Development*, 24, 486-494. DOI: 10.1016/j.cogdev.2009.09.001

Kaufmann, L., & von Aster, M. (2012). The diagnosis and management of dyscalculia. *Deutsches Ärzteblatt International*, 109, 767-778. DOI: 10.3238/arztebl.2012.0767

Kline, P. (2000). *A psychometrics primer*. London, UK: Free Association Books.

Landerl, K., Bevan, A. & Butterworth, B. (2004). Developmental dyscalculia and basic numerical capacities: A study of 8–9-year-old students. *Cognition*, 93, 99-125. DOI: 10.1016/j.cognition.2003.11.004

Logan, G. D., & Cowan, W. B. (1984). On the ability to inhibit thought and action: A theory of an act of control. *Psychological Review*, 91, 295-327. DOI: 10.1037/0033-295X.91.3.295

Lyons, I. M., & Ansari, D. (2015). Numerical order processing in children: From reversing the distance-effect to predicting arithmetic. *Mind, Brain, and Education*, 207-221. DOI: 10.1111/mbe.12094

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

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

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, 714-726. DOI: 10.1111/desc.12152

- Lyons, I. M., Vogel, S. E., & Ansari, D. (2016). On the ordinality of numbers: A review of neural and behavioral studies. *Progress in Brain Research*, 227, 187-221. DOI: 10.1016/bs.pbr.2016.04.010
- Majerus, S., Poncelet, M., Greffe, C., & van der Linden, M. (2006). Relations between vocabulary development and verbal short-term memory: The relative importance of short-term memory for serial order and item information. *Journal of Experimental Child Psychology*, 93, 95-119. DOI: 10.1016/j.jecp.2005.07.005
- Mammarella, I. C., Hill, F., Devine, A., Caviola, S., & Szűcs, D. (2015). Math anxiety and developmental dyscalculia: a study on working memory processes. *Journal of Clinical and Experimental Neuropsychology*, 37, 878-887. DOI: 10.1080/13803395.2015.1066759
- Morsanyi, K., Devine, A., Nobes, A. & Szűcs, D. (2013). The link between logic, mathematics and imagination. Evidence from children with developmental dyscalculia and mathematically gifted children. *Developmental Science*, 16, 542-553. DOI: 10.1111/desc.12048
- 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. *The Quarterly Journal of Experimental Psychology*, 70, 2497-2517. DOI: 10.1080/17470218.2016.1246577
- 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 math. (in press). DOI: 10.1111/desc.12645

Passolunghi, M. C., & Siegel, L. S. (2001). Short-term memory, working memory, and inhibitory control in children with difficulties in arithmetic problem solving. *Journal of Experimental Child Psychology*, 80, 44-57. DOI: 10.1006/jecp.2000.2626

Passolunghi, M. C., & Siegel, L. S. (2004). Working memory and access to numerical information in children with disability in mathematics. *Journal of Experimental Child Psychology*, 88, 348-367. DOI: 10.1016/j.jecp.2004.04.002

Piazza, M., Facoetti, A., Trussardi, A.N., Berteletti, I., Conte, S., Lucangeli, D., et al. (2010). Developmental trajectory of number acuity reveals a severe impairment in developmental dyscalculia. *Cognition*, 116, 33-41. DOI: 10.1016/j.cognition.2010.03.012

Price, G. R., Holloway, I., Räsänen, P., Vesterinen, M., & Ansari, D. (2007). Impaired parietal magnitude processing in developmental dyscalculia. *Current Biology*, 17, R1042-R1043. DOI: 10.1016/j.cub.2007.10.013

Rousselle, L., & Noël, M. P. (2007). Basic numerical skills in children with mathematics learning disabilities: A comparison of symbolic vs non-symbolic number magnitude processing. *Cognition*, 102, 361-395. DOI: 10.1016/j.cognition.2006.01.005

Rubinsten, O., & Henik, A. (2009). Developmental dyscalculia: Heterogeneity might not mean different mechanisms. *Trends in Cognitive Sciences*, 13, 92-99. DOI: 10.1016/j.tics.2008.11.002

Formatted: Font: (Default) Times New Roman, 12 pt

Formatted: Font: (Default) Times New Roman, 12 pt

Formatted: Font: (Default) Times New Roman, 12 pt, Font color: Auto

Rubinsten, O., & Sury, D. (2011). Processing ordinality and quantity: The case of developmental dyscalculia. *PLoS ONE*, 6, e24079. DOI: 10.1371/journal.pone.0024079

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. DOI: doi.org/10.1016/j.cognition.2017.04.007

Sattler, J. M., & Dumont, R. (2004). Assessment of children: WISC-IV and WPPSI-III supplement. San Diego: Jerome M. Sattler, Publisher.

Siegler, R. S. & Opfer, J. E. (2003). The development of numerical estimation evidence for multiple representations of numerical quantity. *Psychological Science, 14*, 237-250.

DOI: 10.1111/1467-9280.02438

Skagerlund, K., & Träff, U. (2014). Development of magnitude processing in children with developmental dyscalculia: space, time, and number. *Frontiers in Psychology, 5*, 675.

Formatted: Font: (Default) Times New Roman, 12 pt

Szűcs, D., Devine, A., Soltész, F., Nobes, A., & Gabriel, F. (2013). Developmental dyscalculia is related to visuo-spatial memory and inhibition impairment. *Cortex, 49*, 2674-2688. DOI: 10.1016/j.cortex.2013.06.007

Szűcs, D., Nobes, A., Devine, A., Gabriel, F., & Gebuis, T. (2013). Visual stimulus parameters seriously compromise the measurement of approximate number system acuity and comparative effects between adults and children. *Frontiers in Psychology, 4*:444. DOI: 10.3389/fpsyg.2013.00444

Swanson, H. L. (2011). Working memory, attention, and mathematical problem solving: A longitudinal study of elementary school children. *Journal of Educational Psychology, 103*, 821-837. DOI: 10.1037/a0025114

Tabachnick, B.G. & Fidell, L.S. (2007). *Using multivariate statistics*. Allyn & Bacon/Pearson Education.

Formatted: Font: (Default) Times New Roman, 12 pt

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

Vicario, C. M., Rappo, G., Pepi, A., Pavan, A., & Martino, D. (2012). Temporal abnormalities in children with developmental dyscalculia. *Developmental Neuropsychology, 37*, 636-652.

Formatted: Font: (Default) Times New Roman, 12 pt, Font color: Auto

- Vincent, D., & Crumpler, M. (2007). *Hodder Group Reading Tests 1–3 (II)*. London: Hodder Education.
- von Aster, M. G., & Shalev, R. S. (2007). Number development and developmental dyscalculia. *Developmental Medicine and Child Neurology*, *49*, 868–873. DOI: 10.1111/j.1469-8749.2007.00868.x
- Vogel, S.E., Remark, A., & Ansari, D. (2015). Differential processing of symbolic numerical magnitude and order in first-grade children. *Journal of Experimental Child Psychology*, *129*, 26–39. DOI: 10.1016/j.jecp.2014.07.010
- 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*, 469-482. DOI: 10.1080/20445911.2017.1282490
- Wechsler, D. (2003). *Wechsler intelligence scale for children* (4th ed.). Oxford, England: Psychological Corp.
- Williams, J. (2005). *Mathematics assessment for learning and teaching*. London: Hodder Education.

**Appendix: Parental Report of Everyday Ordering Ability**

*Please circle the number which you feel best applies to your child for each question.*

My son/daughter:

*Can easily adjust to changes in routine.*

(1=very much disagree; 7=very much agree)

1----2----3----4----5----6----7

*Understands how the calendar works.*

(1=very much disagree; 7=very much agree)

1----2----3----4----5----6----7

*Can easily recall the order in which past events happened.*

(1=very much disagree; 7=very much agree)

1----2----3----4----5----6----7

*Is able to plan a sequence of activities independently.*

(1=very much disagree; 7=very much agree)

1----2----3----4----5----6----7

*Finds it easy to learn new activities which involve a sequence of actions that have to be performed in a particular order (e.g., when learning to play computer or board games).*

(1=very much disagree; 7=very much agree)

1----2----3----4----5----6----7

*Would find it easy to remember a phone number.*

(1=very much disagree; 7=very much agree)

1----2----3----4----5----6----7

*Can organise their own time when doing certain tasks (e.g., can decide in what order to do different pieces of homework).*

(1=very much disagree; 7=very much agree)

1---2---3---4---5---6---7