Cognition 165 (2017) 26–38

Contents lists available at ScienceDirect

Cognition

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

Original Articles

Unpacking symbolic number comparison and its relation with arithmetic in adults

Delphine Sasanguie [a,b,*], Ian M. Lyons [c], Bert De Smedt [d], Bert Reynvoet [a,b]

[a] Brain and Cognition, KU Leuven, 3000 Leuven, Belgium
[b] Faculty of Psychology and Educational Sciences @Kulak, 8500 Kortrijk, Belgium
[c] Department of Psychology, Georgetown University, 20057 Washington D.C., United States
[d] Parenting and Special Education, KU Leuven, 3000 Leuven, Belgium

A R T I C L E I N F O

Article history:
Received 15 July 2016
Revised 13 April 2017
Accepted 24 April 2017
Available online 28 April 2017

Keywords:
Digit comparison
Arithmetic
Working memory
Long-term memory
Order processing

A B S T R A C T

Symbolic number – or digit – comparison has been a central tool in the domain of numerical cognition for decades. More recently, individual differences in performance on this task have been shown to robustly relate to individual differences in more complex math processing – a result that has been replicated across many different age groups. In this study, we 'unpack' the underlying components of digit comparison (i.e. digit identification, digit to number-word matching, digit ordering and general comparison) in a sample of adults. In a first experiment, we showed that digit comparison performance was most strongly related to digit ordering ability – i.e., the ability to judge whether symbolic numbers are in numerical order. Furthermore, path analyses indicated that the relation between digit comparison and arithmetic was partly mediated by digit ordering and fully mediated when non-numerical (letter) ordering was also entered into the model. In a second experiment, we examined whether a general order working memory component could account for the relation between digit comparison and arithmetic. It could not. Instead, results were more consistent with the notion that fluent access and activation of long-term stored associations between numbers explains the relation between arithmetic and both digit comparison and digit ordering tasks.

1. Introduction

Symbolic representations of numbers are often investigated using a digit comparison task, both in children (e.g., Holloway & Ansari, 2009; Mussolin, Meijas, & Noël, 2010; Sasanguie, Göbel, Moll, Smets, & Reynvoet, 2013) and in adults (e.g., Castronovo & Göbel, 2012; Moyer & Landauer, 1967). In this task, participants need to indicate the larger of two presented digits. However, just what this task measures – and hence, indirectly, what underlying processes it indexes – remains somewhat unclear.

When comparing two digits (e.g., 8 and 9), several cognitive skills are required (see also Purpura & Ganley, 2014). First, one has to identify the symbol that one is presented with as an Arabic numeral (i.e., a digit). This skill has previously been investigated by symbol knowledge tasks (i.e., "Is the symbol a digit") or digit identification tasks (i.e., "Associate the digit with the magnitude it represents") and rapid automatized (digit) naming (RAN). Moreover, several studies have demonstrated that this basic skill of identifying symbols as numerals is associated with arithmetic performance (e.g. Cirino, 2011; Koponen, Salmi, Eklund & Aro, 2013; Mazzocco & Grimm, 2013; Purpura, Baroody, & Lonigan, 2013; van der Sluis, de Jong, & van der Leij, 2004; Vanbinst, Ghesquière, & De Smedt, 2012, 2015).

Second, these culturally acquired symbols need to be matched with their phonological counterpart. In cultural letter acquisition and reading, for example, an efficient mapping process between phonological and orthographic elements is crucial (Blomert & Willems, 2010). Similarly (but see McCloskey & Schubert, 2014), at some point in development – though still no consensus exists about when and in which order exactly – digits are mapped to their corresponding verbal number words (e.g. know that '/two/' is equal to '2') (Benoit, Lehalle, Molina, Tijus, Jouen, 2013; Purpura & Ganley, 2014). Recently, researchers have shown that this audiovisual mapping skill was related to arithmetic achievement

* Corresponding author at: Research Unit Brain and Cognition, KU Leuven, Faculty of Psychology and Educational Sciences @Kulak, Etienne Sabbelaan 53; 8500 Kortrijk, Belgium.

E-mail address: Delphine.Sasanguie@kuleuven.be (D. Sasanguie).

http://dx.doi.org/10.1016/j.cognition.2017.04.007
0010-0277/

in both adults (Sasanguie & Reynvoet, 2014) and elementary school children (Lyons, Price, Vaessen, Blomert, & Ansari, 2014).

Third, knowledge about the ordinal relations among or sequence of Arabic numerals is necessary to perform digit comparison (Turconi, Campbell, & Seron, 2006). To decide whether 9 is larger than 8, one requires ordinal information about the digits that goes beyond the simple count list. Here again, researchers recently demonstrated the presence of a relation between numerical order processing and arithmetic (Attout & Majerus, 2015; Lyons & Ansari, 2015; Lyons & Beilock, 2011; Lyons et al., 2014).

Finally, after comparing two digits, a decision must be made about which of the two digits is numerically larger/smaller. This decision process may be numerically specific (pertaining specifically to numerical stimuli), it may be more general (i.e., common to other non-numerical comparisons – e.g., which letter is closer to Z), or some combination thereof (e.g., Holloway & Ansari, 2008).

Furthermore, during the past decade, numerous researchers have demonstrated that performance on digit comparison tasks is concurrently as well as predictively associated with arithmetic. This relation is very robust and has been observed in typically developing children (e.g., Bugden & Ansari, 2011; De Smedt, Verschaffel, & Ghesquière, 2009; Holloway & Ansari, 2009; Kolkman, Kroesbergen, & Leseman, 2013; Lyons et al., 2014; Sasanguie, De Smedt, Defever, & Reynvoet, 2012; Sasanguie et al., 2013; Vanbinst et al., 2015; Vogel, Remark, & Ansari, 2015; for a meta-analysis, see Schneider et al., 2016) and in children with mathematical learning difficulties (Brankaer, Ghesquière, & De Smedt, 2014; De Smedt & Gilmore, 2011; Landerl, Fussenegger, Moll, & Willburger, 2009; Rousselle & Noël, 2007; Vanbinst, Ghesquière, & De Smedt, 2014; for a meta-analysis, see Schwenk et al., 2017). Though this relation appears to be similar in adults as well, it is worth noting that studies on this topic with adults are surprisingly few (Castronovo & Göbel, 2012; Lyons & Beilock, 2011). When performance was measured by means of reaction times (RT), this relation was the most consistent, although performance measures such as accuracy and distance effects (i.e. faster and more accurate responses to digits that are numerically further away; Moyer & Landauer, 1967) have revealed similar results (for a review, see De Smedt, Noël, Gilmore, & Ansari, 2013). In sum, individuals who are better at indicating which of two presented Arabic numerals is numerically larger tend to have better arithmetic scores. Here we examined whether the process or processes that contributed most to explaining digit comparison could also explain some or all of the widely reported relation between digit comparison and arithmetic performance.

In a first experiment, in addition to digit comparison performance itself, we also assessed each of the four candidate processes discussed above (i.e., digit identification, digit to number-word audiovisual matching, digit order judgment and letter comparison). First, we assessed which of these processes captured unique variance in the digit comparison task. Of those that did, we next asked whether they could account for some or all of the relation between digit comparison and arithmetic abilities.

2. Experiment 1

2.1. Method

2.1.1. Participants

Sixty-seven university students participated for monetary compensation. Seven participants were removed from the analyses because of missing data or because they performed too slowly or made too many errors (>3*SD* above the group mean) in one of the experimental tasks. Consequently, the final sample comprised 60 adults (M_{age} = 20.43 years; *SD* = 2.73; 50 females).

2.1.2. Procedure

Participants were tested in groups of about 20, accompanied by two experimenters. First, all participants performed a paper-and-pencil arithmetic test, which was administered in group (i.e. the instructions were read aloud for the whole group and then the participants were requested to fill in their own page). Next, the subjects performed the experimental computer tasks measuring the candidate cognitive processes discussed above: (1) fast identification, (2) audiovisual matching, (3) order judgment and (4) comparison. For this, participants sat together in the same room, but could work individually on their own computer screen, at their own pace. Each of these tasks was presented in a numerical and a non-numerical condition, leading to eight tasks which were conducted in a randomized order using a Latin square design. Afterwards, participants individually performed two reading tests.

All experimental tasks (see Fig. 1) were conducted using a 15-inch color screen connected to a computer running the Windows 7 operating system. Stimulus presentation and recording of the behavioral data (reaction times and error rates) were controlled by E-prime Professional software, version 2.0 (Psychological Software Tools, Pittsburgh, PA, USA). In all tasks, each trial was preceded with a fixation cross of 600 ms, after which two stimuli appeared (one on the left and one on the right side of the screen) and remained on the screen for 1000 ms. Afterwards, a blank was presented until a response was detected. Participants could respond (by pressing 'a' for indicating the left or 'p' for indicating the right stimulus on an AZERTY keyboard) during the stimulus presentation or during the blank. The visual stimuli were presented in white against a black background (courier new font, 40pt, Bold). The same accounted for the audiovisual tasks, except that in these tasks, an auditory presented stimulus was presented simultaneously with the two visual stimuli. These auditory stimuli (i.e., verbal number words or letter speech sounds) were digitally recorded (sampling rate 44.1 kHz, 16-bit quantization) by a Dutch female speaker. Recordings were band-pass filtered (180–10.000 Hz), resampled at 22.05 kHz, and matched for loudness. The sounds were presented binaurally through loudspeakers at about 65 dB SPL. On all tasks, subjects were instructed to respond as quickly and as accurately as possible. The inter-trial interval was 1500 ms. Each task started with five practice trials in which feedback was provided. During the experimental trials, there was no feedback.

2.1.3. Measures

2.1.3.1. Experimental tasks.

2.1.3.1.1. Fast identification tasks. The stimulus set consisted of single digits (2–8[1]; numerical condition) or letters (I, J, N, L, N, O, P, R, T, U; letter condition) and randomly chosen symbols $, #, £, @, § and €. Participants needed to indicate which of two presented stimuli was the digit (numerical condition) or the letter (non-numerical condition), in contrast to the random symbol. The digits and the letters were equally presented on the left (n = 7) and on the right side of the screen (n = 7), resulting in 14 trials, each presented 5 times. This way, the trial list of the digit and the letter condition each consisted of 70 experimental trials.

2.1.3.1.2. Audiovisual matching tasks. An auditory stimulus was presented, i.e., a number word (e.g., [aχt] (eight)), in the numerical condition or a letter-speech sound (e.g., [i.] (i)), in the non-numerical letter condition, together with two visually presented

[1] Similar as in the first study investigating the potential role of order-related processing in number comparison (Turconi et al., 2006), we used digits 2–8 in order to create close/sequential (distance 1) pairs (2–3/3–4 and 6–7/7–8) and far (distance 3) pairs (2–5/3–6 and 4–7/5–8). To keep all tasks similar, this number range was also used in the other numerical tasks (i.e., fast digit identification and audiovisual matching).

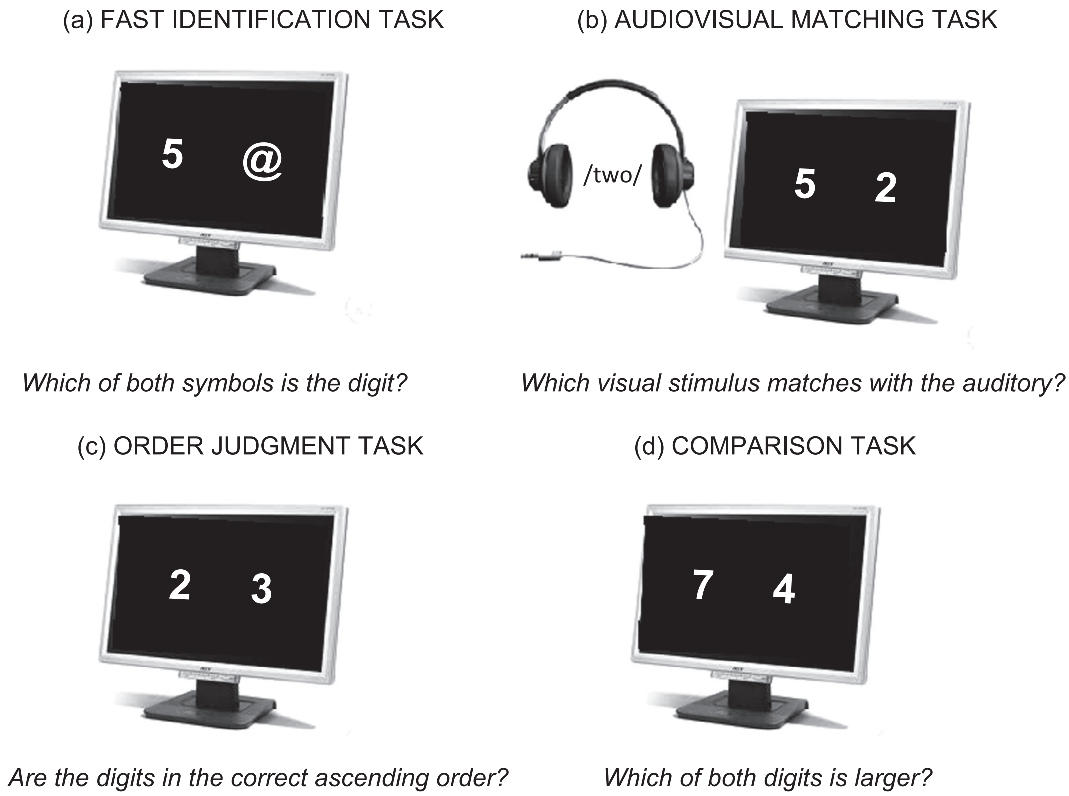

Fig. 1. Illustration of the four experimental tasks (in the numerical condition) of Experiment 1: (a) fast identification task, (b) audiovisual matching task, (c) order judgment task and (d) comparison task.

stimuli – the same digits and letters as in the identification tasks. Participants were instructed to choose the right or left visually presented stimulus that matched the auditory stimulus. The auditory stimulus matched as frequently with the left as with the right stimulus (each stimulus was presented equally often on the left and on the right). Moreover, distance was manipulated between the two visually presented stimuli, resulting in two conditions of each 16 trials – presented twice: trials with distance 'one' (e.g., 2–3 or I-J) and trials with distance 'three' (e.g., 2–5 or O-R). The final trial list thus consisted of 64 experimental digit trials and 64 experimental letter trials.

2.1.3.1.3. Order judgment tasks. The same digits and letters were used as in the previous tasks. To keep the task design similar for all experimental tasks, two digits or two letters appeared on the screen and participants indicated whether these stimuli were presented in left-to-right ascending (digit condition) or alphabetical (letter condition) order ('a') or not ('p'). Similarly as in the audiovisual matching task, stimuli were counterbalanced between left and right sides of the screen (i.e., each combination of stimuli was presented in an ascending and a descending order) and distance was manipulated, resulting in two conditions (distance one and distance three) of eight trials each. These eight trials were presented 4 times, so that the final trials list consisted of 64 digit and 64 letter trials.

2.1.3.1.4. Comparison tasks. Trial lists were exactly the same as in the order judgment tasks, except that the participants now were instructed to choose which of two presented stimuli was larger (numerical condition) or closest to 'Z' (non-numerical letter condition).

2.1.3.2. Standardized tests.

2.1.3.2.1. Arithmetic. Participants were tested with the Tempo Test Arithmetic (Tempo Test Rekenen - TTR; De Vos, 1992). This timed paper- and- pencil test consists of five subtests: one for each type of operation (addition, subtraction, multiplication and division) and one with mixed operations. Forty items of increasing complexity were presented in each subtest, and participants were given one minute to solve as many problems as possible. One point was awarded for each correct item. Participants scored on average 150.88 (SD = 23.74) out of 200.

2.1.3.2.2. Word reading. The One-Minute Reading test (Een-Minuut-Test - EMT; Brus & Voeten, 1991) was administered. The EMT is a standardized test used to examine the accuracy and the speed (fluency) of word reading. The test consists of 116 unrelated words, divided in four rows with increasing complexity. The participant is asked to read the words as quickly and accurately as possible, and is stopped after one minute. The score is the total number of correct words read in one minute. Participants scored on average 96.23 (SD = 13.21) out of 116.

2.1.3.2.3. Non-word reading. The Klepel (van den Bos, Spelberg, Scheepstra, & de Vries, 1994) is a standardized reading test used to examine the accuracy and speed of non-word reading. The test consists of 116 non-words (e.g., taaf, worenpuis, krobbelon, gepaltnomeng etc.), divided in four rows with increasing complexity. The words are constructed such that they follow the phonotatic structure of the Dutch language. The participant is asked to read the non-words as quickly and accurately as possible, and is stopped after two minutes. The score is the total number of correct non-words read in two minutes. Participants scored on average 101.35 (SD = 14.30) out of 116.

2.2. Results

Data of the experimental identification task and audiovisual matching task in the non-numerical condition (i.e., with letters and letter speech sounds) as well as the data from the two reading tasks were not analyzed here as they bear no direct relation to the theoretical question at hand (i.e., understanding digit comparison and how/why this task relates to Arithmetic scores). These additional tasks and tests were included here as the dataset as a whole was intended to ask several questions, including the basic cognitive underpinnings of Reading ability. However, that is beyond the scope of the paper presented here.

2.2.1. Overall performance

The mean accuracies and median reaction times are displayed in Table 1. In most tasks, accuracy performance was at ceiling. An analysis of the association between accuracy and speed revealed that, except for the digit order judgment task ($r(58) = -0.01$, $p = 0.91$) and the digit comparison task ($r(58) = 0.17$, $p = 0.18$), participants responded more slowly to be more accurate (e.g., $r(58) = 0.44$, $p < 0.0001$ for the letter order judgment task; $r(58) = 0.21$, $p = 0.10$ for the digit identification task; $r(58) = 0.25$, $p = 0.06$ for the digit audio-visual matching task and $r(58) = 0.25$, $p = 0.06$ for the letter comparison task). To account for this speed-accuracy trade-off (SAT), we used adjusted reaction times (RT), i.e. RT/ACC – also referred to as 'inverse efficiency'-, in all further analyses. This way we could combine reaction times and accuracies: Reaction times remained unchanged with 100% accuracy and increased in proportion with the number of errors (for a similar procedure, see also Simon et al., 2008).

In line with previous research (e.g., Turconi et al., 2006), we analyzed the ascending (e.g., 2–3) and descending trials (e.g., 3–2) in the order judgment tasks separately to examine whether this task really addressed ordinal processing (i.e. deciding whether stimuli are presented in the ascending/alphabetical left-to-right order or not), and participants did not rely on for instance comparison strategies (i.e. deciding which one is larger) to solve this task. Indeed, according to Turconi et al. (2006), the presence of a reversed distance effect (RDE; i.e., faster and more accurate responses to digits that are numerically closer) in the ascending order trials implies order-specific processes such as serial search or direct recognition of order. We observed an RDE in the ascending trials of the letter order judgment task, $F(1, 59) = 7.23$, $p = 0.009$, $\eta_p^2 = 0.109$: trials with distance one were processed 465 ms faster than trials with distance three. In the numerical digit order judgment task, this RDE was not present (we observed a standard/canonical distance effect in the ascending trials, $F(1, 59) = 18.69$, $p < 0.001$, $\eta_p^2 = 0.241$, showing slower responses to digits that are numerically closer), but ascending trials were processed 61 ms faster than the descending trials, $t(59) = -6.09$, $p < 0.0001$, a phenomenon that was not present in the comparison task (the RT difference between ascending and descending trials in that case was only 5 ms, $t < 1$).

2.2.2. Relations with digit comparison

We first investigated which cognitive components underpin digit comparison performance. We did this by asking which of the four candidate predictors (digit-identification, digit-audiovisual-matching, digit-ordering and letter-comparison) were related to digit comparison. Table 2 shows the zero-order correlations between digit comparison and the four candidate predictors, as well as the associations among the four predictors themselves.

Table 2
Zero-order correlations among the experimental tasks of Experiment 1.

		1	2	3	4	5
1	Arithmetic	1				
2	Digit identification	−0.01	1			
3	Digit audiovisual matching	−0.04	0.43**	1		
4	Digit order judgment	−0.44**	0.20	0.26*	1	
5	Digit comparison	−0.39**	0.41**	0.33**	0.51**	1
6	Letter comparison	−0.37**	0.08	0.14	0.42**	0.35**

* $p < 0.05$.
** $p < 0.01$.

All candidate predictors tasks were significantly related to digit comparison (see Table 2).

We next asked which of these predictors was *uniquely* related to digit comparison. We did this via a multiple regression analysis with digit comparison as dependent variable. The significant correlations between predictors imposed no problem for multicollinearity, as all VIF-values were close to 1 (Field, 2009). Results are shown in Table 3, from which it is clear that only digit identification and digit ordering uniquely contributed to the variance in the performance on the digit comparison task. Together, these analyses suggest that individuals who are more proficient at comparing two digits (1) more rapidly identify the symbol that they are presented with as an Arabic numeral, and (2) more rapidly process the ordinal relations among sequences of the Arabic numerals.

2.2.3. Accounting for the relation between digit comparison and arithmetic

After identifying the critical cognitive processes involved in digit comparison, we investigated if these processes explained the well-known and widely replicated relation between number comparison and more complex math abilities (i.e., Arithmetic). Specifically, we identified digit identification and digit ordering as unique predictors of digit comparison performance, so those two tasks were tipped in this section as potential mediators of the relation between digit-comparison and Arithmetic performance. Hereto, we ran a mediation analysis using the Preacher and Hayes (2008) SPSS 'MEDIATE' macro for simultaneously testing multiple mediators within a single analysis. These analyses were bootstrapped, which makes them more accurate with small sample sizes. Moreover, bootstrapping is robust to violations of the normality assumption, which is often violated in smaller samples (MacKinnon, Lockwood, Hoffman, West, & Sheets, 2002; Preacher & Hayes, 2008). A mediation analysis uncovers whether there is a significant *indirect effect* (quantified as the product of the unstandardized path coefficients, ab) of the mediator(s) that accounts for some portion of the *total effect* (written as c) observed between the predictor and the outcome variable. The remaining (unmediated) *direct effect* of the original predictor on the outcome variable is denoted c'. In this framework, the model is constrained by the assumption that $c = ab + c'$. Unlike in a standard multiple regression analysis, here it is explicitly asked what portion of the relation between digit comparison performance and arithmetic that can be accounted for by the mediating variables. When ab is significant but c' is not, this is interpreted as *full* mediation; when both ab

Table 1
Mean accuracies, median reaction times (and the corresponding standard deviations), per notation, of the four experimental tasks of Experiment 1.

	Fast identification	Audiovisual matching	Order judgment		Comparison	
	Digits	*Digits*	*Digits*	*Letters*	*Digits*	*Letters*
Accuracies (proportion)	0.99 (0.01)	0.96 (0.03)	0.94 (0.04)	0.80 (0.12)	0.96 (0.03)	0.84 (0.10)
Reaction times (msec)	387 (31)	544 (44)	616 (83)	1713 (613)	478 (46)	1163 (353)

Table 3
Multiple regression analysis with the three digit tasks and letter comparison as predictors and digit comparison as dependent variable (Experiment 1).

Independent variables	Standardized β	t	p
Digit identification	0.291	2.487	0.016
Digit audiovisual matching	0.091	0.768	0.446
Digit order judgment	0.359	3.006	0.004
Letter comparison	0.163	1.407	0.165

$F(4,55) = 8.851$, $p < 0.001$, $R^2 = 0.391$.

and c' remain significant, this is interpreted as *partial* mediation (Preacher & Hayes, 2008). Because we used the bootstrapping method recommended by Preacher and Hayes (2008), our conclusions are based on the confidence intervals. Strictly for reference for those more familiar with traditional methods, classical hypothesis testing indicators (i.e., p-values) are also included. However, the reader is strongly encouraged to consider confidence intervals as the primary mode of assessing statistical validity - in particular whether confidence intervals include 0.

Fig. 2 displays the mediation results for Experiment 1. Arithmetic was our outcome and the performance on the digit comparison task served as predictor (Fig. 2a). The performance on the other cognitive tasks that uniquely contributed to digit comparison in the regression analysis above (i.e., digit identification and digit ordering) were included in the model as mediator variables (Fig. 2b). Only digit ordering performance uniquely explained (mediated) a significant portion of the relation between number comparison and arithmetic: bootstrap point estimate $a_1b_1 = -0.09$; $SE = 0.04$; *95% CI* $= -0.176$ to -0.013; $p = 0.037$. The mediating effect of digit identification was not significant: bootstrap point estimate $a_2b_2 = 0.04$; $SE = 0.03$; *95% CI* $= -0.010$ to 0.110; $p = 0.206$. Note also that the overall mediation effect was only *partial* (i.e., the total effect c was not completely accounted for by the mediators), indicating that there remains an aspect of the variance in digit comparison performance which is uniquely captured by comparing relative symbolic magnitudes (cf. the significant direct effect c'). In sum, this mediation analysis demonstrated that especially knowledge about the ordinal sequence of Arabic numerals that goes beyond the counting list plays an impor-

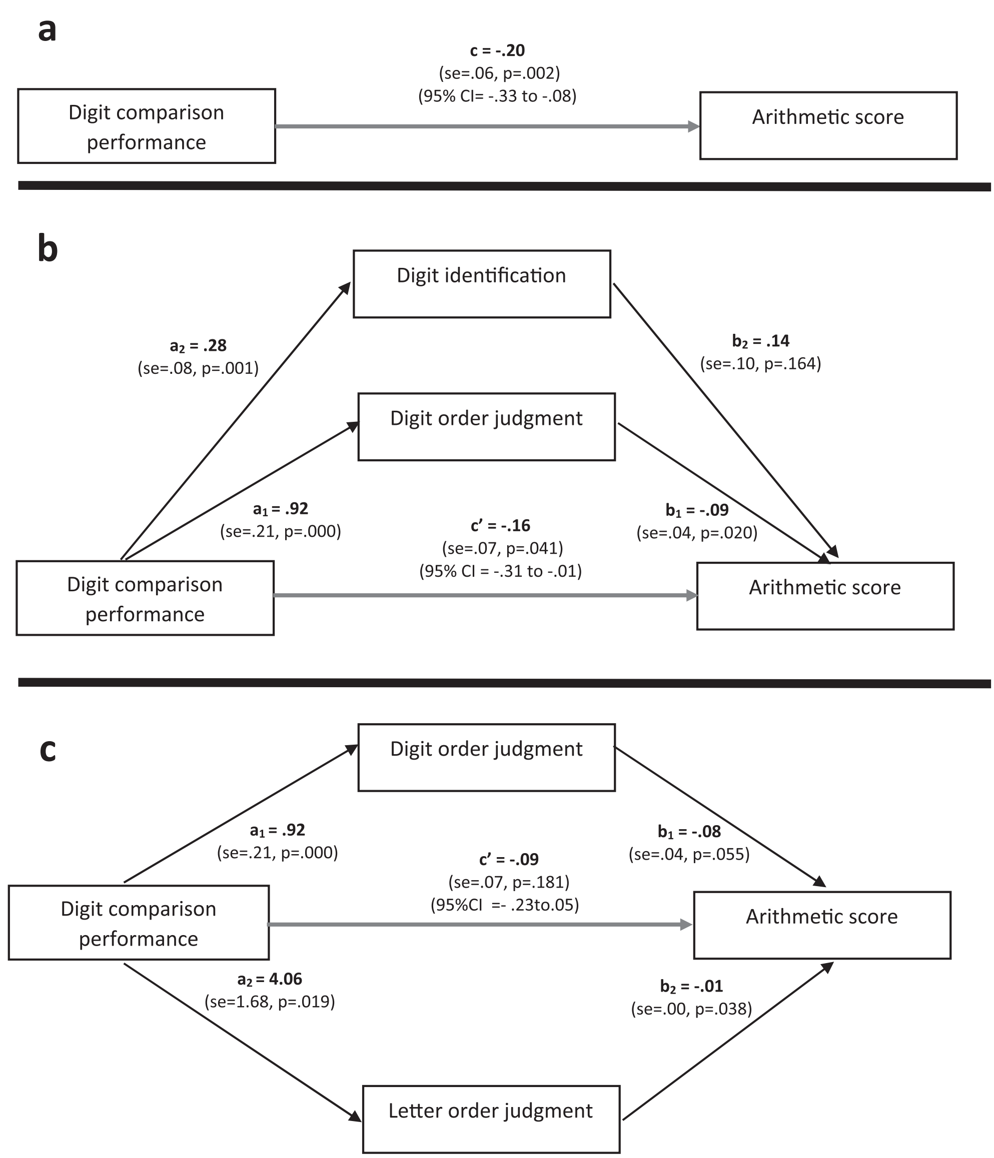

Fig. 2. Mediation models of Experiment 1. Fig. 2**a** demonstrates the total effect c of theoriginal predictor (digit comparison performance) on the outcome variable (arithmetic score), after taking into account the mediating variables. Fig. 2**b** displays the results of the mediation effects of digit order judgment (a1 * b1) and digit identification (a2 * b2), as well as the unmediated direct effect c'. Fig. 2**c** shows the mediating effect of both digit order judgment (a1 * b1) and letter order judgment (a2 * b2) on the relation between digit comparison and arithmetic.

tant role in explaining part of the relation between digit comparison and Arithmetic.

Since this study concerns cross-sectional data, we additionally tested whether the mediation was also present in the opposite direction (i.e., whether the relation between order judgment and arithmetic is mediated by digit comparison). Results of this model showed a significant total effect, path coefficient $c = -0.13$; $SE = 0.03$; $95\%\ CI = -0.193$ to -0.058; $p < 0.001$ and a significant direct effect, path coefficient $c' = -0.09$; $SE = 0.04$; $95\%\ CI = -0.170$ to -0.016; $p = 0.19$, but no significant effect of the mediator variable digit comparison, bootstrap point estimate $ab = 0.03$; $SE = 0.02$; $95\%\ CI = -0.092$ to 0.006; $p = 0.124$. In other words, while digit ordering explained a significant portion of the relation between digit comparison and arithmetic, digit comparison did not explain a significant portion of the relation between digit ordering and arithmetic.

2.2.4. Understanding the mediating role of digit ordering

In this section, we assessed whether the mediation effect of digit ordering demonstrated in the previous section was due to *numerically-specific* or *more general ordinal processing* factors. We performed an additional mediation analysis with digit comparison as the initial predictor and Arithmetic as the outcome variable, but with both digit order judgment and letter order judgment as simultaneous mediator variables. Fig. 2c demonstrates that both tasks uniquely accounted for the relation between digit comparison and arithmetic performance. For digit ordering, bootstrap point estimate $a_1b_1 = -0.07$; $SE = 0.04$; $95\%\ CI = -0.162$ to -0.005; $p = 0.078$. For letter ordering, a_2b_2 -0.04; $SE = 0.03$; $95\%\ CI = -0.107$ to -0.0001; $p = 0.127$. Interestingly, both order tasks together *fully* accounted for the relation between digit comparison performance and Arithmetic (see Fig. 2c in which the c' path is no longer significant).

2.3. Discussion

Of the four candidate cognitive processes (i.e., digit identification, digit to number-word audiovisual matching, digit ordering and general comparison), only digit identification and digit ordering captured unique variance in digit comparison performance. This result suggests that the ability to rapidly identify a visually presented symbol as a digit and the ability to assess the ordinality of a sequence of Arabic digits are key skills that may underpin how one compares two digits in terms of their relative magnitude. Especially the former is actually not surprising, since visual recognition is a prerequisite to perform comparison: one cannot compare digits if one does not visually recognizes the symbols he/she is confronted with. In contrast, matching a symbol with its phonological counterpart (i.e., as measured with the audiovisual matching task) and general comparison processes (i.e., as measured with the letter comparison task) are not predictive for digit comparison performance. Since audiovisual matching did not contribute to the variance in digit comparison performance and previous studies at least have demonstrated that the performance on the audiovisual matching task contributes to the variance in arithmetic achievement in both adults (Sasanguie & Reynvoet, 2014) and elementary school children (Lyons et al., 2014), it could possibly be assumed that the contribution of audiovisual matching to arithmetic performance is unique and cannot be explained by cognitive processes that are shared by audiovisual matching and digit comparison. On the other hand, in the current study, there was no contribution of audiovisual matching to the variance in arithmetic performance and the audiovisual matching task used in the current study and the one mentioned in the study of Sasanguie and Reynvoet (2014) differ in instructions and task design. Sasanguie and Reynvoet (2014) asked participants to decide whether the auditory presented number word and the visually presented digit were the same or different. In contrast, to keep the task design as similar as possible in the different tasks used in the current study, participants here had to decide if the number word they heard corresponded to the left or right presented digit. Consequently, the use of two visually presented digits instead of one and the different instructions might account for the contrasting results observed in the current study (i.e., no relation between digit-number word matching and arithmetic) versus the ones described in the study of Sasanguie and Reynvoet (2014) (i.e., a correlation of -0.36, $p < 0.01$ between digit-number word matching performance and arithmetic). Moreover, these task differences imply that it is to date also not clear yet whether the contribution of audiovisual matching to arithmetic is unique or can be explained by overlapping cognitive processes shared with digit comparison.

Furthermore, we observed a relation between digit comparison performance and arithmetic in adults, which replicates a result that has been shown many times in studies with children, but surprisingly few in studies with adults (Castronovo & Göbel, 2012; Lyons & Beilock, 2011). When entering the two key cognitive processes of digit comparison discussed above into a mediation analysis, we observed that only digit ordering could account for a significant portion of the relation between digit comparison performance and arithmetic. Importantly, when examining the reversed mediation model (i.e., order judgment as predictor, arithmetic as outcome and comparison as mediator variable), results demonstrated that comparison did not mediate the relation between digit order judgment and arithmetic. This latter result is important because it helps to constrain one's interpretation of the underlying mechanism: digit ordering can explain the relation between digit comparison and arithmetic, but the reverse is not true (comparison cannot explain the relation between digit ordering and arithmetic). Broadly, this result is in line with several recent studies which provided evidence for an important role of order in digit processing and arithmetic (Lyons & Beilock, 2011; Lyons et al., 2014).

On the other hand, it is important to note that this was only a partial mediation (Fig. 2b). Moreover, even for the subset of variance accounted for by the digit ordering task, it remains somewhat unclear precisely by what mechanism digit ordering explains (in part) the relation between digit comparison and arithmetic. In particular, is this result due to numerically specific or more general processes? Crucially, therefore, our results also show that both explanations hold a degree of validity. Both digit and letter ordering were unique mediators of the relation between digit comparison and arithmetic. Importantly, this indicates that the digit ordering task captured unique variance that could not be accounted for by the non-numerical letter ordering task. Moreover, the letter ordering task also captured unique variance of its own, indicating a role for a more general (i.e., non-numerical) ordering component.

Together, for the first time, the results of Experiment 1 thus show that ordinal processing is key to explaining the widely reported relation between digit comparison performance and arithmetic (e.g. Brankaer et al., 2014; Bugden & Ansari, 2011; Castronovo & Göbel, 2012; De Smedt & Gilmore, 2011; De Smedt et al., 2009; Holloway & Ansari, 2009; Kolkman et al., 2013; Landerl et al., 2009; Lyons & Beilock, 2011; Lyons et al., 2014; Rousselle & Noël, 2007; Sasanguie et al., 2012, 2013; Vanbinst et al., 2014; Vanbinst et al., 2015; Vogel et al., 2015). There are however two limitations to what we can infer from Experiment 1. First, in Experiment 1, pairs of stimuli (digits or letters) were used in the order judgment tasks, and recent work indicates that using pairs in such an order judgment task may in fact tap a combination of comparison and ordering processes. Indeed, Vogel et al. (2015) used an order judgment task with pairs and found a standard/canonical distance effect (i.e., faster and more accurate responses to digits that are numerically further away) in first graders, whereas by contrast, Lyons and Ansari (2015) used three digits (triplets) and observed a robust reversed distance effect (RDE; faster and more accurate responses to digits that are numerically closer) in first graders. In Experiment 1 of the current study, an RDE was found only for the letter order judgment task. Thus, one possible explanation for the explanatory role of digit ordering is that it is simply a slightly modified version of the digit comparison task. Hence, what is needed is a digit ordering task that is thought to more reliably activate ordinal processing – i.e., a digit order judgment task with triplets.

Second, the letter order judgment task in Experiment 1 did elicit a large RDE, indicating it tapped ordinal processing. However, it is not clear what this task measures. It seems plausible to surmise it is tapping domain general ordinal processing, such as serial order working memory. On the other hand, it would be more ideal to have a task that tests this proposition more directly.

Therefore, we conducted a second experiment. In Experiment 2, our aims were (1) to replace the pairs version of the digit order judgment task with a triplet version, and (2) to replace the letter order judgment task with a serial order working memory task. In this way, we could more directly test the notions (1) that it is in fact ordinal processing that explains the relationship between digit comparison and arithmetic, and (2) whether the critical ordinal processing component is specific to number-processing or indicative of more general ordinal processing.

3. Experiment 2

3.1. Method

3.1.1. Participants

Seventy-nine university students participated for monetary compensation. Six participants were removed from the analyses because they performed too slowly or made too many errors (>3*SD* above the group mean) in one of the experimental tasks. Consequently the final sample comprised 73 adults (M_{age} = 19.87 years; *SD* = 5.79; 59 females).

3.1.2. Procedure

Participants were tested individually or in small groups, accompanied by an experimenter. All participants first performed an arithmetic test, followed by four experimental tasks (administered in a randomized order using a Latin square design): (1) a digit order judgment task with pairs (included for replication purposes); (2) a digit order judgment task with triplets; (3) a serial order working memory task; and (4) a digit comparison task. All experimental tasks were conducted using a 15-in. color screen connected to a computer running the Windows 7 operating system. Stimulus presentation and recording of the behavioral data (reaction times and error rates) were controlled by E-prime Professional software, version 2.0 (Psychological Software Tools, Pittsburgh, PA, USA). In all tasks, stimuli were presented in white against a black background (Courier New Font 40, Bold) and subjects were instructed to respond as quickly and as accurately as possible. Each task started with five practice trials in which feedback was provided. During the experimental trials, there was no feedback.

3.1.3. Measures

The digit order judgment task with pairs, the digit comparison task and the arithmetic test were identical as in Experiment 1. Therefore, we only describe the tasks that were added in Experiment 2: the digit order judgment task with triplets and the serial order working memory task (see also Fig. 3).

3.1.3.1. Digit order judgment task with triplets. Each trial started with a fixation cross presented for 600 ms. Afterwards, three single digits (range: 1–9) were horizontally presented on the screen for 1000 ms, followed by a blank screen. Participants were instructed to decide whether the digits were presented in an order (either ascending or descending). If all three digits were presented in an order, participants were instructed to press the 'a' key on an AZERTY keyboard with their left hand. If the sequence of digits was not presented in an order (e.g., 1-3-2), participants were instructed to press the 'p' key with their right hand. They could respond during stimulus presentation or during the blank screen. The inter-trial interval was 1500 ms. In total, 84 trials were administered. In half of the trials, the triplets had a small distance (span between the smallest and largest digit was 3) and in the other half of the trials triplets had a large distance (span between the smallest and largest digit could vary from 5 to 9). This resulted in 6 conditions of 14 randomly presented trials each, namely small ascending trials (e.g., 1-2-3), small descending trials (e.g., 3-2-1), small non-order trials (e.g., 3-1-2), large ascending trials (e.g., 1-3-5), large descending trials (e.g., 5-3-1) and large non-order trials (e.g., 5-1-3).

3.1.3.2. Serial order working memory (OWM) task. This task was based on the serial order short- term memory task used in

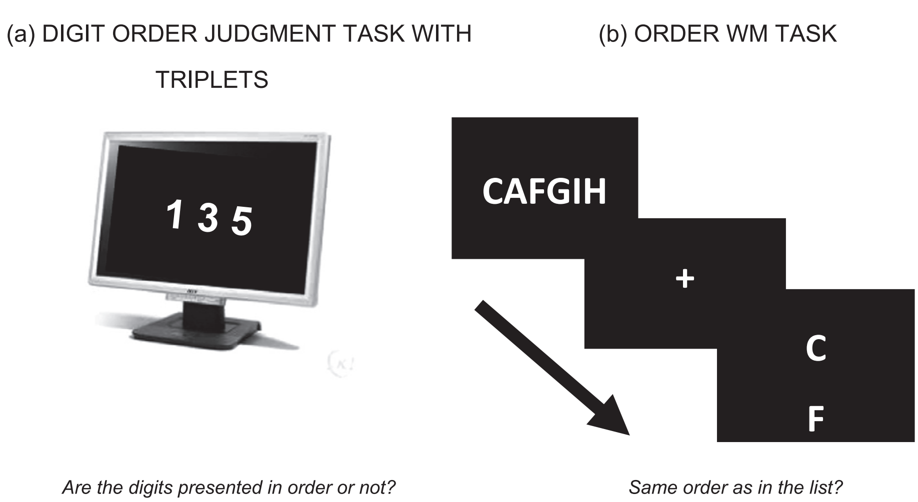

Fig. 3. Illustration of (a) the digit order judgment task with triplets and (b) the serial order working memory task (OWM) used in Experiment 2.

Attout, Fias, Salmon, and Majerus (2014). The stimulus set consisted of letters from A to I. First, for 2500 ms, participants were presented with a list of six letters ordered horizontally on the screen (e.g., 'I, H, A, F, D, C') and were instructed to try to remember the list as accurately as possible (i.e., the encoding phase). Afterwards, a fixation cross was displayed for a variable duration (random Gaussian distribution centered on a mean duration of 4500 ± 1500 ms), during which the participants had to keep the order of the list in mind (i.e., the maintenance phase). Finally, the retrieval phase consisted of two vertically ordered letter stimuli (to eliminate the possibility that the task could be completed by mere visuospatial matching) on which the participants had to decide within 3000 ms whether the letter presented on the top of the screen had occurred in a more leftward position in the memory list than the letter presented on the bottom of the screen (by pressing 'a' and 'p' for 'yes' and 'no' respectively on an AZERTY keyboard). The inter-trial interval was 1500 ms. In the retrieval phase, the positional distance between the letters as they were presented in the memory list varied from 2 to 5, resulting in 4 conditions of 12 trials each (i.e., 48 trials in total). In half of the trials, the answer was correct, in the other half, the answer was incorrect. The alphabetical distance between the letters was kept constant (distance of 3).

3.2. Results

3.2.1. Overall performance

Participants scored on average 139.99 (SD = 19.91) out of 200 on the Arithmetic test. The mean accuracies and median reaction times on the experimental tasks are displayed in Table 4. Similar to Experiment 1, we used adjusted reaction times (RT), i.e. RT/ACC – also referred to as 'inverse efficiency' – in all further analyses.

Also similar to Experiment 1, we analyzed the ascending and descending trials of the order judgment tasks separately to examine whether a reversed distance effect (RDE; faster and more accurate responses to digits that are numerically closer) in the ascending trials was present and the tasks thus actually tapped into ordinal processing. As in Experiment 1, in the digit order judgment task with pairs, this RDE was not present (we again observed a standard/canonical distance effect in the ascending trials, $F(1, 72) = 39.78$, $p < 0.001$, $\eta_p^2 = 0.356$). By contrast, in the digit order judgment task with triplets, a clear RDE was present in the ascending trials, $F(1, 72) = 26.087$, $p < 0.001$, $\eta_p^2 = 0.266$, and in the descending trials, $F(1, 72) = 6.270$, $p = 0.015$, $\eta_p^2 = 0.080$, indicating that trials with a small distance were processed faster than trials with a large distance (354 ms faster in the ascending and 339 ms faster in the descending condition).

3.2.2. Relations with digit comparison

We again first investigated the relation between the candidate predictors and digit comparison. Table 5 shows the zero-order correlations between digit comparison and the three other experimental tasks included in Experiment 2 (the two order judgments tasks and the OWM task), as well as among the three experimental tasks themselves. All three experimental tasks were significantly related to digit comparison (see Table 5).

Table 5
Zero-order correlations among the experimental tasks of Experiment 2.

		1	2	3	4
1	Arithmetic	1			
2	Digit order pairs	−0.32**	1		
3	Digit order triplets	−0.51**	0.57**	1	
4	OWM task	−0.10	0.31**	0.46**	1
5	Digit comparison	−0.31**	0.50**	0.47**	0.27*

Note. OWM = Order working memory task.
* $p < 0.05$.
** $p < 0.01$.

Table 6
Multiple regression analysis with the two order judgment tasks and the OWM task as predictors and digit comparison as dependent variable (Experiment 2).

Independent variables	Standardized β	t	p
Digit order pairs	0.334	2.714	0.008
Digit order triplets	0.258	1.957	0.054
OWM task	0.047	0.409	0.684

$F(3,69) = 9.900$, $p < 0.001$, $R^2 = 0.301$.

Next, we examined which of these predictors was *uniquely* related to digit comparison via a multiple regression analysis with digit comparison as dependent variable. The significant correlations between predictors imposed no problem for multicollinearity, as all VIF-values were close to 1 (Field, 2009). Results are shown in Table 6, from which it is clear that only the two order judgment tasks uniquely contributed to the variance in the performance on the digit comparison task. Note however that the contribution of the order judgment task with triplets was only marginally significant ($p = 0.054$).

3.2.3. Candidate Mediators: Digit ordering with pairs versus triplets and the role of serial order working memory

To test whether (1) it is in fact ordinal processing that explains the relation between digit comparison and arithmetic, and (2) whether this critical ordinal processing component is specific to number processing or indicative of more general ordinal processing, we next conducted four mediation analyses (see Fig. 4) using the same Preacher and Hayes (2008) bootstrapping method as in Experiment 1. Again, the reader is strongly encouraged to consider confidence intervals as the primary mode of assessing statistical validity – in particular whether confidence intervals include 0.

For replication purposes, we first conducted the exact same mediation analysis as in Experiment 1, with digit comparison performance as predictor, arithmetic as outcome and digit order judgment task with pairs as mediator (see Fig. 4a). Similar to Experiment 1, we observed a significant mediating effect of the digit order judgment task with pairs, bootstrap point estimate $ab = -0.04$; $SE = 0.02$; *95% CI* = −0.08 to −0.01; $p = 0.11$. Note that, in contrast to the partial mediation obtained with this task in Experiment 1, here we observed a full mediation effect of the digit order judgment task with pairs.

Second, we replaced the performance on the digit order judgment task with pairs by the performance on the digit order judgment task with triplets in the abovementioned mediation analysis (see Fig. 4b). Again, we observed a full mediation effect,

Table 4
Mean accuracies, median reaction times (and the corresponding standard deviations) of the four experimental tasks of Experiment 2.

	Digit order pairs	Digit order triplets	OWM task	Digit comparison
Accuracies (proportion)	0.94 (0.04)	0.88 (0.08)	0.84 (0.10)	0.96 (0.03)
Reaction times (msec)	587 (97)	978 (244)	1535 (272)	463 (64)

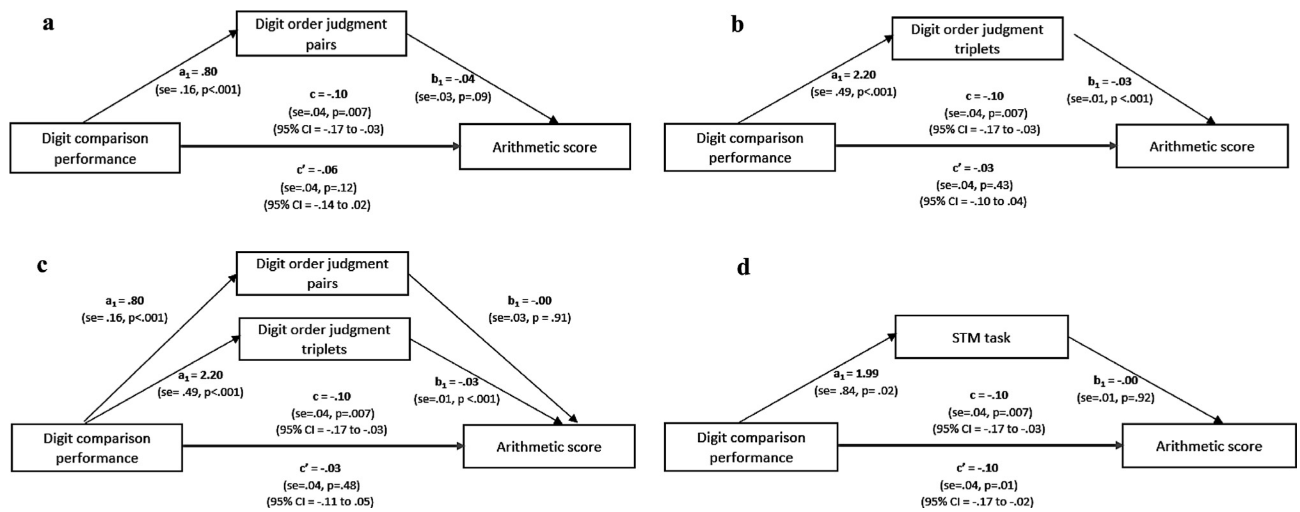

Fig. 4. Mediation models of Experiment 2. In all mediation models, the performance on the digit comparison task is the predictor and arithmetic scores are the outcome, only the mediator variable changes throughout the 4 models. Fig. 4**a** demonstrates a full mediation effect of the digit order judgment task with pairs. Fig. 4**b** displays a full mediation effect of the digit order judgment task with triplets. Fig. 4**c** shows that when entering both order tasks into the model, only the digit order judgment task with triplets is a significant mediator. Fig. 4**d** demonstrates that the performance on the OWM task does not mediate the relation between digit comparison performance and arithmetic outcome.

bootstrap point estimate $ab = -0.07$; $SE = 0.02$; *95% CI* $= -0.12$ to -0.03; $p = 0.003$.

Third, because the regression analysis (Table 6) indicated that both ordering tasks uniquely contributed to the variance in digit comparison performance, we ran the same mediation analysis again, but now included both ordering tasks simultaneously (see Fig. 4c). We observed a full mediation because the total effect c was fully accounted for by the significant indirect effect of the digit order judgment task with triplets, bootstrap point estimate $a_2b_2 = -0.07$; $SE = 0.03$; *95% CI* $= -0.12$ to -0.03; $p = 0.006$. However, the indirect effect of the digit order task with pairs was not significant, bootstrap point estimate $a_1b_1 = -0.003$; $SE = 0.02$; *95% CI* $= -0.04$ to 0.04; $p = 0.907$.

Finally, to examine whether the observed critical mediating effect of ordinal processing is specific to number or reflects more general ordinal processing, we replaced the performance on the order judgment task(s) by the performance on the serial order working memory (OWM) task as mediator in the model (see Fig. 4d). In contrast to all previous mediation models of Experiment 2, serial order working memory did not mediate the relation between digit comparison performance and arithmetic, bootstrap point estimate $ab = -0.00$; SE = 0.01; 95% CI = -0.02 to 0.02; p = 0.93.

3.2.4. Understanding the mediating role of digit ordering versus the non-mediating role of serial order working memory

In this section, we investigate in more depth why we observed a significant mediating role for numerical ordering but not for domain-general non-numerical ordering (Fig. 4) by unpacking the mediation models. Specifically, we examine the a and b paths of the different models separately.

The models displayed in Fig. 4a and b demonstrated that numerical (i.e. digit) ordering ability play a crucial role in explaining the well-established relation between digit comparison and arithmetic, irrespective of which version of the order judgment task (pairs or triplets) was used. Crucially, however, the model in Fig. 4c showed that when entering both tasks simultaneously into the mediation model, only the ordering task with triplets – the version of the ordering task thought to more reliably activate ordinal processing – was significant. To unravel this mediation in more detail, we broke down the a path and the b path of the mediation model displayed in Fig. 4c. With regard to the a paths, we conducted two partial correlation analyses to investigate (1) whether the significant relation between *comparison* and order pairs holds when controlling for the performance on the order task with triplets; and (2) whether the relation between *comparison* and the performance on the order task with triplets holds when controlling for the performance on the order task with pairs. Results showed that both relationships remained significant. Comparison ~ Pairs: $r_p(70) = 0.31$, $p = 0.007$; Comparison ~ Triplets: $r_p(70) = 0.26$, $p = 0.026$. Both versions of the order judgment task provide a unique contribution to the performance on the digit comparison task. With regard to the b paths, we conducted two partial correlation analyses to investigate (1) whether the significant relation between *arithmetic* and order pairs holds when controlling for the performance on the order task with triplets; and (2) whether the relation between *arithmetic* and the performance on the order task with triplets holds when controlling for the performance on the order task with pairs. Results showed that the relationship between arithmetic and order pairs disappeared and only the relationship between arithmetic and order triplets remained significant. Arithmetic ~ Pairs: $r_p(70) = -0.04$, $p = 0.720$; Arithmetic ~ Triplets: $r_p(70) = -0.42$, $p < 0.001$. Together, these results thus suggest that, although both versions of the order judgment task uniquely contribute to digit comparison performance, it is only the performance on the order judgment task with triplets that relates uniquely with arithmetic. Thus, combining the a and b paths, as is the case in a mediation analysis, we see that only ordering triplets is able to uniquely mediate the relation between digit comparison performance and arithmetic.

Fig. 4d showed that the performance on the serial order working memory (OWM) task did not mediate the relation between digit comparison performance and arithmetic. With regard to the a path, performance on the OWM task did indeed relate to performance on the digit comparison task: $r(71) = 0.27$, $p = 0.021$. However, for the b path, from Table 5 one can see that, even at the zero-order level, performance on the OWM task was unrelated to arithmetic scores, $r(71) = -0.10$, $p = 0.425$. Hence, when combining the a and b paths, we see that OWM cannot explain (mediate) the relation between digit comparison and arithmetic. Instead, as discussed above, our results suggest, we need to call upon the more domain-specific aspects of ordinal processing, as is thought to be indexed by the order judgment tasks (and strongest still by the triplets version thereof, Fig. 4c).

3.3. General discussion

The most commonly used task to investigate symbolic magnitude representations is the digit comparison task (i.e., deciding which of two presented digits is numerically larger). What this task measures exactly and which underlying cognitive processes are tapped when performing this task, however, remain unclear to date. Moreover, in many previous studies (mostly with children and a few with adults), a concurrent as well as a predictive relation between digit comparison performance and individual differences in arithmetic has been reported. To our knowledge, the current study is the first to examine which of four possible candidate cognitive processes (digit identification, digit to number-word audiovisual matching, digit ordering, and general comparison) are uniquely related to digit comparison performance. Moreover, it is unclear which (if any) of these processes may account for some or all of the widely reported relation between digit comparison and arithmetic abilities in adults.

In Experiment 1, we observed that numerical ordering (measured with an order judgment task with pairs of digits) is crucial in explaining the largely reported relation between digit comparison performance and individual differences in arithmetic. Moreover, next to this numerically-specific contribution, we also observed a mediating role for the performance on a non-numerical (i.e., letter) order judgment task. In Experiment 2, we added another version of the numerical order task to ensure that we really tapped ordinal processing and we examined more directly whether the mediating role of the letter-order judgment task was indeed due to a domain-general ordinal process, such as serial order working memory. Results from Experiment 2 (1) confirmed that performance on the numerical order task was not just a significant mediator because it shared a comparison component with the comparison task and (2) indicated that serial order working memory could not explain the relation between digit comparison performance and arithmetic. We elaborate on these two findings below.

First, results of Experiment 2 showed that both the performance on the order judgment task with pairs and on the order judgment tasks with triplets were significant full mediators of the relation between digit comparison performance and arithmetic. In contrast to Experiment 1, the order judgment task with pairs was not a partial but a full mediator here. Most likely, this difference is however not qualitatively meaningful, since the difference between the *p*-values of the direct paths *c'* in the two experiments is negligible ($p = 0.04$ in Experiment 1 versus $p = 0.12$ in Experiment 2). The similar findings obtained with the two versions of the ordering task might at first glance suggest similarity in processing, but there are also critical differences. Table 5 for instance demonstrated that the correlation between the two versions of the task is $r(71) = 0.57$, $p < 0.001$, which can be considered a moderate correlation. If both tasks measured the exact same underlying cognitive process, a much higher correlation might be expected. Moreover, in line with previous studies, we observed a standard/canonical distance effect for the order judgment task with pairs (e.g., see also Turconi, Jemel, Rossion, & Seron, 2004; Vogel et al., 2015); in contrast, we saw a reversed distance effect in the ordering task with triplets (e.g., see also Lyons & Ansari, 2015). Finally, when entering both versions of the ordering task simultaneously into the mediation model, performance on the ordering task with triplets turned out to be the only significant mediator. Methodologically, the above suggests that both task versions may index different aspects of numerical processing. The absence of the RDE in the pairs task for instance could be due to variability in strategies to solve this task: a comparison strategy on some trials and an order judgment on others (resulting in faster processing of the ascending trials compared to the descending trials, a phenomenon that is not present in a regular comparison task). Therefore, the pairs version of the task possibly captures a combination of comparison and ordinal processing. In contrast, the robust RDE observed in the ordering task with triplets more reliably suggests the activation of ordinal processing (Franklin, Jonides, & Smith, 2009; Lyons & Beilock, 2011, 2013).

Hence, theoretically, with respect to the relation between digit comparison and arithmetic, our results suggest that, although digit comparison performance comprises both a comparison and an ordinal processing component, only the ordinal processing component seems to fully account for the relation between the performance on the digit comparison task and individual differences in arithmetic. The latter is important in the light of the emerging interest of numerical cognition researchers in the difference between the cardinality of symbolic numbers (i.e., the number of items in a set that a symbol/digit represents – commonly measured using a number comparison task) and their ordinality (i.e., the sequencing of number symbols, e.g. five is the fifth number – it comes after four and before six; Lyons & Beilock, 2013; for a review, see Lyons, Vogel, & Ansari, 2016) and their respective contributions to arithmetic skills. Especially in adults, there is a lack of research addressing the relationship between those three components. To the best of our knowledge, only Goffin and Ansari (2016) have investigated in adults whether the different mechanisms that appear to underlie cardinal and ordinal processing of symbolic numbers related differently to more complex mathematical skills. These researchers observed that the reversed distance effect (RDE) on the ordinality task and the standard numerical distance effect (NDE) from the cardinality task (i.e., the two hallmarks of the two respective tasks) were not significantly correlated with one another, suggesting that the mechanisms underlying cardinal and ordinal processing are distinct in adults. Although the difference between the two constructs is not the primary focus of the current study, our observations (for instance in Table 6) confirm that cardinality and ordinality are most probably not reducible to one and the same mechanism. Rather, our results in fact suggest that the performance on the digit comparison task does not purely reflect cardinal, but also contains a distinct ordinal component. With respect to arithmetic fluency, Goffin and Ansari (2016) found that both cardinality and ordinality accounted for unique variance in math achievement. The latter is however not supported by our results. Indeed, the observations in both Experiment 1 and Experiment 2 that ordinality processing fully accounted for the relation between digit comparison and arithmetic suggest that 'cardinality' (i.e., performance on the digit comparison task) no longer explains unique variance in individual differences in arithmetic achievement once ordinality processing is taken into account. It must be noted however that in the study of Goffin and Ansari (2016) distance effects were used in contrast to the performance measure (i.e., inverse efficiency, a combination of reaction time and accuracies) in the current study, which might account for the different results. Finally, one other study (Lyons & Beilock, 2011) investigated the relationship between cardinality, ordinality and arithmetic, but in that study cardinality was measured with a non-symbolic comparison task (i.e., deciding which of two dot arrays contains more dots). Similar to the current study with symbolic comparison, Lyons and Beilock (2011) found that ordinal processing fully mediated the relationship between a non-symbolic comparison task and arithmetic. Together, these findings thus provide evidence of an important role for ordering abilities – *over and above cardinality processing* – in adult mathematical skills, an observation which in fact nicely squares with the learning trajectory outlined in the early numeracy educator's guide by Frye et al. (2013), namely that fluency with mental number comparisons builds on fluency with number-after relations and provides a basis for mental arithmetic.

Second, results of Experiment 2 showed that domain-general ordinal processing (serial order working memory) *does not* explain the relation between digit comparison and arithmetic. This suggest that it may be long-term ordinal associations as opposed to short-term ordinal processing that are key to explaining the relation between digit comparison and arithmetic. Several theoretical models describe how order information is processed and, grossly, they can be divided into two categories: 'associative chaining models' and 'position marker models' (Abrahamse, Van Dijck, Majerus, & Fias, 2014). Associative chaining models state that people form inter-item associations between adjacent items (e.g., between 3 and 4) and each item serves as a trigger for the next (or previous) item in the sequence (e.g., 3 triggers 4; Serra & Nairne, 2000). According to position marker models, people connect each item to a specific position (e.g., begin vs. end items) and these connections are recalled when the sequence of the items has to be retrieved (Solway, Murdock, & Kahana, 2012). For instance, to remember a list of animal names (e.g., dog – cat – bird), a first item (e.g., dog) is marked by code "1", a second item (e.g., cat) to code "2", a third item (e.g., bird) by code "3" et cetera. More recently, some authors propose that a combination of both models is possible (e.g., Abrahamse et al., 2014; Caplan, 2015; Solway et al., 2012). Abrahamse et al. (2014) suggest that position markers are used as a basic mechanism to store and retrieve order information, whereas chaining may only be used in particular situations. For instance, position markers of a sequence of items that had to be retained for a longer time could become connected with each other and may act like chaining mechanisms (Abrahamse et al., 2014). The latter is what we believe is occurring in the (numerical) ordinal judgment task – and is consequently what may be crucial for (learning) arithmetic. Indeed, if the storage, retrieval and processing of digits is considered, it is clear that adults are already highly familiar with the counting sequence (Lyons & Beilock, 2009, 2013; Nieder, 2009). When processing digits, there is not only a connection between the symbols and the quantities they represent, but also a connection between the symbols themselves, beyond their link with the underlying quantities (Abrahamse et al., 2014; Lyons, Ansari, & Beilock, 2012; Lyons & Beilock, 2009; Nieder, 2009). Therefore, it might be the case that, to solve the numerical order judgment task, adults simply have to activate these associations between the digits themselves from their long-term memory (i.e., chaining based mechanisms). This hypothesis is in line with Lyons and Beilock (2013), who suggested that making ordinal judgments about digits relies on retrieval processes.

Whether these long-term ordinal associations, which appear to be key in explaining individual differences in arithmetic, are specific to numbers was not directly addressed in the current study. However, since Experiment 1 showed that next to digit ordering also letter ordering significantly mediates the relation between digit comparison and arithmetic and Experiment 2 showed that this mediation process cannot be explained on the basis of serial order working memory involvement, it is plausible to assume that also the letter order judgment task reflects those long-term ordinal associations. If this would be true, one could expect a full mediation of the relation comparison – arithmetic by letter ordering performance, similar as is the case for numerical ordering in Experiment 2. To test this hypothesis more directly, we post hoc conducted a mediation analysis on the data of Experiment 1, with digit comparison as predictor, arithmetic as outcome and letter order judgment performance as the only mediator. We observed a partial mediation, as the total effect c (bootstrap point estimate = −0.20; *SE* = 0.06; *95% CI* = −0.33 to −0.08; $p = 0.002$) was not totally accounted for by the indirect effect of letter ordering (bootstrap point estimate = −0.05; *SE* = 0.03; *95% CI* = −0.11 to −0.008; $p = 0.091$). The direct effect c′ was still significant, bootstrap point estimate = −0.15; *SE* = 0.06; *95% CI* = −0.28 to −0.03; $p = 0.017$. These data suggest that the long-term ordinal associations which appeared to be key for arithmetic are not domain-general, but most probably number-specific. A possible explanation for the presence of a partial mediation of letter ordering could be that in contrast to the counting sequence of digits that is most probably stored via chaining mechanisms in long-term memory, adults are likely less familiar with the order of letters – and certainly in case of the letters used in Experiment 1 (I, J, N, L, N, O, P, R, T, U), which are not the first ones of the alphabet (Gevers, Reynvoet, & Fias, 2003; Jou & Aldridge, 1999; Turconi et al., 2004). Turning back to the theoretical models of order processing, it is therefore likely that in case of letter order processing, a combination of position marker strategies and chaining strategies is used (for a similar suggestion, see Caplan, 2015). For example, if we have to decide whether K is before or after another letter, we first decide on the position of K (i.e., the position marker) and then do chaining. So possibly, in order to solve a letter order judgment task, we create different position markers, most probably based on how we learn the alphabet (and dependent on our short-term memory capacity; Jou & Aldridge, 1999): ABCDEFG . . .HIJKLMNOP . . ..QRSTUVW. . .XYZ. If we have to decide whether R is before/after U: we first activate the position marker 'Q' and then do chaining. It could therefore be that this combination of strategies present in the letter order judgment task only evoked a partial mediation of the relation comparison-arithmetic, since it is especially the associative chaining that is responsible for the mediation. Of course, this is speculative and should be investigated further. If, on the other hand, letters would (similar to digits) be stored in long-term memory based on associative chaining instead of a combination of position marker strategies and chaining strategies, the process of fluently accessing and activating the ordinal associations in long-term memory in the case of letters could still only evoke a partial mediation effect of the relation between comparison and arithmetic, because letter processing per se is not predictive for/involved in arithmetic, only the (possibly domain-general) associative chaining process is.

4. Conclusion

Symbolic magnitude representations or the 'cardinality' of symbols is commonly investigated using a digit comparison task. Moreover, performance on this task is associated with individual differences in arithmetic scores, both in children and adults. In two experiments, this study showed first that (a) an important aspect of digit comparison performance is ordinal processing, and (b) this ordinal processing perhaps reflects more associative chaining mechanisms in long-term memory, rather than general ordinal processing such as serial order working memory. Second, the associative chaining mechanism in long-term memory may be central to unpacking the consistently reported relation between digit comparison and more complex numerical skills, such as arithmetic. Future studies should unravel, however, whether this associative chaining mechanism only accounts for numbers, or whether also a fluent access and activation of non-numerical chaining plays a role in explaining individual differences in arithmetic.

Acknowledgements

Delphine Sasanguie is a postdoctoral research fellow of the Research Foundation Flanders (www.fwo.be). This research was also supported by the Research Fund KU Leuven and the study was approved by the local ethics committee (S56006 and G-2015

09 317). The authors thank Evy Mentens, Laura Sourbron, Jasmien Heyns en Marijke Vanderperren for their help with the data collection.

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.cognition.2017.04.007.

References

Abrahamse, E., Van Dijck, J. P., Majerus, S., & Fias, W. (2014). Finding the answer in space: The mental whiteboard hypothesis on serial order in working memory. *Frontiers in Human Neuroscience, 8*. http://dx.doi.org/10.3389/tnhum.2014.00932.

Attout, L., Fias, W., Salmon, E., & Majerus, S. (2014). Common neural substrates for ordinal representation in short-term memory, numerical and alphabetical cognition. *PLoS ONE, 9*(3), e92049. http://dx.doi.org/10.1371/journal.pone.0092049.

Attout, L., & Majerus, S. (2015). Working memory deficits in developmental dyscalculia: The importance of serial order. *Child Neuropsychology: A Journal on Normal and Abnormal Development in Childhood and Adolescence, 21*(4), 432–450.

Benoit, L., Lehalle, H., Molina, M., Tijus, C., & Jouen, F. (2013). Young children's mapping between analogical, verbal and digital codes. *Cognition, 129*, 95–101.

Blomert, L., & Willems, G. (2010). Is there a causal link from a phonological awareness deficit to reading failure in children at familial risk for dyslexia? *Dyslexia, 16*, 300–317.

Brankaer, C., Ghesquière, P., & De Smedt, B. (2014). Numerical magnitude processing deficits in children with mathematical difficulties are independent of intelligence. *Research in Developmental Disabilities, 35*, 2603–2613.

Brus, B. Th., & Voeten, M. J. M. (1991). *EMT: Een-minuut-test*. Amsterdam, The Netherlands: Berkhout Nijmegen B.V.

Bugden, S., & Ansari, D. (2011). Individual differences in children's mathematical competence are related to the intentional but not automatic processing of Arabic numerals. *Cognition, 118*, 35–47.

Caplan, J. B. (2015). Order-memory and association-memory. *Canadian Journal of Experimental Psychology, 69*(3), 221–232.

Castronovo, J., & Göbel, S. M. (2012). Impact of high mathematics education on the number sense. *PLoS One, 7*, e33832. http://dx.doi.org/10.1371/journal.pone.0033832.

Cirino, P. T. (2011). The interrelationships of mathematical precursors in kindergarten. *Journal of Experimental Child Psychology, 108*, 713–733.

De Smedt, B., & Gilmore, M. (2011). Defective number module or impaired access? Numerical magnitude processing in first graders with mathematical difficulties. *Journal of Experimental Child Psychology, 108*, 278–292.

De Smedt, B., Noël, M. P., Gilmore, C., & Ansari, D. (2013). How do symbolic and non-symbolic numerical magnitude processing skills relate to individual differences in children's mathematical skills? A review of evidence from brain and behavior. *Trends in Neuroscience and Education, 2*, 48–55.

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.

De Vos, T. (1992). *TTR: Tempo test rekenen*. Amsterdam, The Netherlands: Berkhout Nijmegen B.V.

Field, A. (2009). *Discovering statistics using SPPS* (3rd ed.). London: Sage Publications.

Franklin, M. S., Jonides, J., & Smith, E. E. (2009). Processing of order information for numbers and months. *Memory & Cognition, 37*, 644–654.

Frye, D., Baroody, A. J., Burchinal, M., Carver, S. M., Jordan, N. C., & McDowell, J. (2013). Teaching math to young children: A practice guide (NCEE 2014–4005). Washington, DC: National Center for Education Evaluation and Regional Assistance (NCEE), Institute of Education Sciences, U.S. Department of Education. Retrieved from the NCEE website: <http://whatworks.ed.gov>.

Gevers, W., Reynvoet, B., & Fias, W. (2003). The mental representation of ordinal sequences is spatially organized. *Cognition, 87*, B87–B95.

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.

Holloway, I. D., & Ansari, D. (2008). Domain-specific and domain-general changes in children's development of number comparison. *Developmental Science, 11*, 644–649.

Holloway, I. D., & Ansari, D. (2009). Mapping numerical magnitudes onto symbols: The numerical distance effect and individual differences in children's mathematics achievement. *Journal of Experimental Child Psychology, 103*, 17–29.

Jou, J., & Aldridge, J. W. (1999). Memory representation of alphabetic position and interval information. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 25*, 680–701.

Kolkman, M. E., Kroesbergen, E. H., & Leseman, P. P. M. (2013). Early numerical development and the role of non-symbolic and symbolic skills. *Learning and Instruction, 25*, 95–103.

Koponen, T., Salmi, P., Eklund, K., & Aro, T. (2013). Counting and RAN: Predictors of arithmetic calculation and reading fluency. *Journal of Educational Psychology, 105*, 162–175.

Landerl, K., Fussenegger, B., Moll, K., & Willburger, E. (2009). Dyslexia and dyscalculia: Two learning disorders with different cognitive profiles. *Journal of Experimental Child Psychology, 103*, 309–324.

Lyons, I. M., & Ansari, D. (2015). Numerical order processing in children: From reversing the distance-effect to predicting arithmetic. *Mind, Brain and Education, 9*, 207–221.

Lyons, I. M., Ansari, D., & Beilock, S. L. (2012). Symbolic estrangement: Evidence against a strong association between numerical symbols and the quantities they represent. *Journal of Experimental Psychology: General, 141*(4), 635–641.

Lyons, I. M., & Beilock, S. L. (2009). Beyond quantity: Individual differences in working memory and the ordinal understanding of numerical symbols. *Cognition, 113*(2), 189–204.

Lyons, I. M., & Beilock, S. L. (2011). Numerical ordering ability mediates the relation between number-sense and mathematical competence. *Cognition, 121*, 256–261.

Lyons, I. M., & Beilock, S. L. (2013). Ordinality and the nature of symbolic numbers. *The Journal of Neuroscience, 33*, 17052–17061.

Lyons, I. M., Vogel, S., & Ansari, D. (2016). On the Ordinality of Numbers: A review of neural and behavioral studies. *Progress in Brain Research, 227*, 187–221.

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

MacKinnon, D. P., Lockwood, C. M., Hoffman, J. M., West, S. G., & Sheets, V. (2002). A comparison of methods to test mediation and other intervening variable effects. *Psychological Methods, 7*, 83–104.

Mazzocco, M. M. M., & Grimm, K. J. (2013). Growth in rapid automatized naming from Grades K to 8 in children with math or reading disabilities. *Journal of Learning Disabilities, 46*, 517–533.

McCloskey, M., & Schubert, T. (2014). Shared versus separate processes for letter and digit identification. *Cognitive Neuropsychology, 31*, 437–460.

Moyer, R. S., & Landauer, T. K. (1967). Time required for judgments of numerical inequality. *Nature, 215*, 1519–1520.

Mussolin, C., Meijas, S., & Noël, M.-P. (2010). Symbolic and nonsymbolic number comparison in children with and without dyscalculia. *Cognition, 115*, 10–25.

Nieder, A. (2009). Prefrontal cortex and the evolution of symbolic reference. *Current Opinion in Neurobiology, 19*, 99–108.

Preacher, K. J., & Hayes, A. F. (2008). Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behavior Research Methods, 40*, 879–891.

Purpura, D. J., Baroody, A. J., & Lonigan, C. J. (2013). The transition from informal to formal mathematical knowledge: Mediation by numeral knowledge. *Journal of Educational Psychology, 105*, 453–464.

Purpura, D. J., & Ganley, C. M. (2014). Working memory and language: Skill specific or domain general relations to mathematics? *Journal of Experimental Child Psychology, 122*, 104–121.

Rousselle, L., & Noël, M.-P. (2012). Basic numerical skills in children with mathematics learning disabilities: A comparison ofsymbolic vs. non-symbolic number magnitude processing. *Cognition, 102*, 361–395.

Sasanguie, D., De Smedt, B., Defever, E., & Reynvoet, B. (2012). Association between basic numerical abilities and mathematics achievement. *British Journal of Developmental Psychology, 30*, 344–357.

Sasanguie, D., Göbel, S. M., Moll, K., Smets, K., & Reynvoet, B. (2013). Approximate number sense, symbolic number processing, or number–space mappings: What underlies mathematics achievement? *Journal of Experimental Child Psychology, 114*, 418–431.

Sasanguie, D., & Reynvoet, B. (2014). Adults' arithmetic builds on fast and automatic processing of arabic digits: Evidence from an audiovisual matching paradigm. *PLoS ONE, 9*(2), e87739. http://dx.doi.org/10.1371/journal.pone.0087739.

Schneider, M., Beeres, K., Coban, L., Simon, M., Schmidt, S. S., Stricker, J., & De Smedt, B. (2016). Associations of non-symbolic and symbolic numerical magnitude processing with mathematical competence: A meta-analysis. *Developmental Science*. http://dx.doi.org/10.1111/desc.12372.

Schwenk, C., Sasanguie, D., Kuhn, J.-T., Kempe, S., Doebler, P., & Holling, H. (2017). (Non)symbolic magnitude processing in mathematically low-achieving and dyscalculic children: A meta-analysis. *Res Dev Disabil, 64*, 152–167.

Serra, M., & Nairne, J. S. (2000). Part-set cuing of order information: Implications for associative theories of serial order memory. *Memory & Cognition, 28*, 847–855.

Simon, T. J., Takarae, Y., DeBoer, T., McDonald-McGinn, D. M., Zackai, E. H., & Ross, J. L. (2008). Overlapping numerical cognition impairments in children with chromosome 22q11. 2 Deletion or Turner syndromes. *Neuropsychologia, 46*, 82–94.

Solway, A., Murdock, B. B., & Kahana, M. J. (2012). Positional and temporal clustering in serial order memory. *Memory & Cognition, 40*, 177–190.

Turconi, E., Campbell, J. I. D., & Seron, X. (2006). Numerical order and quantity processing in number comparison. *Cognition, 98*, 273–285.

Turconi, E., Jemel, B., Rossion, B., & Seron, X. (2004). Electrophysiological evidence for differential processing of numerical quantity and order in humans. *Cognitive Brain Research, 21*, 22–38.

van den Bos, K. P., Spelberg, H. C. lutje, Scheepstra, A. J. M., & de Vries, J. R. (1994). *De klepel*. Lisse, The Netherlands: Swets & Zeitlinger.

van der Sluis, S., de Jong, P. F., & van der Leij, A. (2004). Inhibition and shifting in children with learning deficits in arithmetic and reading. *Journal of Experimental Child Psychology, 87*, 239–266.

Vanbinst, K., Ghesquière, P., & De Smedt, B. (2012). Numerical magnitude representations and individual differences in children's arithmetic strategy use. *Mind, Brain and Education, 6*, 129–136.

Vanbinst, K., Ghesquière, P., & De Smedt, B. (2014). Arithmetic strategy development and its domain-specific and domain-general cognitive correlates: A longitudinal study in children with persistent mathematical learning difficulties. *Research in Developmental Disabilities, 35*, 3001–3013.

Vanbinst, K., Ghesquière, P., & De Smedt, B. (2015). Does numerical processing uniquely predict first graders' future development of single-digit arithmetic? *Learning & Individual Differences, 37*, 153–160.

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.