



**QUEEN'S
UNIVERSITY
BELFAST**

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.

Morsanyi, K., O'Mahoney, E., & McCormack, T. (2016). Number comparison and number ordering as predictors of arithmetic performance in adults: Exploring the link between the two skills, and investigating the question of domain-specificity. DOI: 10.1080/17470218.2016.1246577

Published in:

The Quarterly Journal of Experimental Psychology

Document Version:

Peer reviewed version

Queen's University Belfast - Research Portal:

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

Publisher rights

© 2016 The Experimental Psychology Society.

This is an Accepted Manuscript of an article published by Taylor & Francis in The Quarterly Journal of Experimental Psychology on 08/11/2016 available online: <http://www.tandfonline.com/doi/full/10.1080/17470218.2016.1246577>

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.

Title: *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*

Running head: *Ordering and arithmetic skills*

Kinga Morsanyi, Eileen O'Mahony & Teresa McCormack

School of Psychology, Queen's University Belfast

Accepted for publication in *The Quarterly Journal of Experimental Psychology* on
26/09/2016.

Corresponding author: Kinga Morsanyi

School of Psychology, Queen's University Belfast, Belfast, BT7 1NN, Northern Ireland, UK
Telephone: +44 (0)28 9097 4326; Fax: +44 (0) 28 9097 4524; email: k.morsanyi@qub.ac.uk

Abstract

Recent evidence has highlighted the important role that number ordering skills play in arithmetic abilities (e.g., Lyons & Beilock, 2011). In fact, Lyons et al. (2014) demonstrated that although at the start of formal mathematics education number comparison skills are the best predictors of arithmetic performance, from around the age of 10, number ordering skills become the strongest numerical predictors of arithmetic abilities. In the current study, we demonstrated that number comparison and ordering skills were both significantly related to arithmetic performance in adults, and the effect size was greater in the case of ordering skills. Additionally, we found that the effect of number comparison skills on arithmetic performance was partially mediated by number ordering skills. Moreover, performance on comparison and ordering tasks involving the months of the year was also strongly correlated with arithmetic skills, and participants displayed similar (canonical or reverse) distance effects on the comparison and ordering tasks involving months as when the tasks included numbers. This suggests that the processes responsible for the link between comparison and ordering skills and arithmetic performance are not specific to the domain of numbers. Finally, a factor analysis indicated that performance on comparison and ordering tasks loaded on a factor which included performance on a number line task and self-reported spatial thinking styles. These results substantially extend previous research on the role of order processing abilities in mental arithmetic.

Key words: arithmetic skills; distance effect; mental number line; number comparison, number ordering.

Recent findings have highlighted the importance of order processing abilities in mathematics, and, in particular, arithmetic skills. However, research in this area is still relatively limited, and, as we describe below, several important questions remain to be answered. This study addressed some of these questions.

One of the first studies that focussed on the role of order processing skills in maths was that of Lyons and Beilock (2011). These authors investigated the relationship between the sensitivity of the “approximate number system” (ANS; e.g., Piazza, Facoetti, Trussardi, Berteletti, Conte & Lucangeli, 2010) (using the dot comparison task), number ordering ability (the ability to quickly decide if numerals appear in the right order), and arithmetic skills in adults. In line with some earlier studies (e.g., Desoete, Ceulemans, De Weerd, & Pieters, 2012; Halberda, Mazocco, & Feigenson, 2008; Piazza et al., 2010), Lyons and Beilock found a significant correlation between non-symbolic comparison performance and maths skills. Nevertheless, the most striking finding of this study was that there was an extremely strong relationship between participants’ ability to process numerical order information and their arithmetic skills (i.e., better number ordering skills were related to better arithmetic performance; $r(52) = -.70, p < .001$). Additionally, number ordering ability mediated the relationship between non-symbolic comparison and mental arithmetic. Lyons and Beilock (2011) argued that number ordering skill are crucial for performing symbolic number operations because, unlike approximate, non-symbolic representations, ordinal representations of numbers are equally precise at any level of magnitude. Thus, these representations are essential for mature mathematical knowledge.

Recent developmental studies have also highlighted the importance of order processing skills in mathematics. Lyons, Price, Vaessen, Blomert, and Ansari (2014) found that whereas number comparison performance (i.e., participants’ ability to quickly decide which of two numerals represents a larger quantity) was a better predictor of early symbolic number skills

than numerical order processing, from around grade 6, numerical order processing was the best predictor of children's arithmetic performance (as compared to seven other numerical predictors of arithmetic skills). There is also newly emerging evidence for an order processing deficit in both adults (Rubinsten & Sury, 2011) and children with developmental dyscalculia (Attout & Majerus, 2014). Finally, Attout, Noel and Majerus (2014) demonstrated that memory for order information was a robust predictor of the development of symbolic number skills in young children. Nevertheless, so far there has been relatively little research into the role of order processing in maths, and existing studies have provided conflicting responses to some basic questions.

In particular, two controversial questions emerge from this literature, both concerning the specificity of the processes involved in number comparison and number ordering. First, to what extent do number comparison and number ordering draw on the same underlying processes? Although there is a strong correlation between performance on these two tasks, and distance effects have been observed in both types of tasks (Goffin & Ansari, 2016), there are also reasons to believe that different underlying processes are involved. For example, Goffin and Ansari (2016) reported that distance effects across the two tasks were not associated with each other and that performance on each task independently predicted unique variance on a test of arithmetic ability. Second, to what extent are the cognitive processes involved in these tasks specific to numbers or are domain general? We will discuss each of these issues in turn.

Common processes in number comparison and number ordering

Comparison tasks including symbolic or non-symbolic numerical stimuli typically show canonical distance effects¹ (i.e., it is harder to discriminate between quantities that are similar to each other, Dehaene, Dupoux, & Mehler, 1990; Moyer & Landauer, 1967). By contrast, several studies (Franklin, Jonides & Smith, 2009; Goffin & Ansari, 2016; Lyons & Beilock, 2013; Turconi, Campbell & Seron, 2006) reported reverse distance effects in the case of tasks where participants had to make judgments about the order of items, suggesting that performance on ordering and comparison tasks relies on different strategies.

Additionally, Lyons and Beilock (2013) presented neural evidence that showed that ordinal and comparison judgments for symbolic numbers do not involve the same areas of the brain. Finally, there is evidence for differential associations between numerical comparison and order judgments and calculation abilities in children (Attout et al., 2014; Lyons et al., 2014; Vogel, Remark, & Ansari, 2015) and adults (Goffin & Ansari, 2016). Interestingly (and in contrast with the findings of Lyons & Beilock, 2013), Franklin et al. (2009) did not find either canonical or reverse distance effects in the case of numerical order judgments that did not involve crossing a decade boundary (but they did find reverse distance effects in the case of boundary-crossing number triads). Franklin et al. (2009) discussed the possibility that, in the case of triads within a decade, canonical and reverse distance effects might both be present, but these effects cancel out each other. Another explanation that Franklin et al. (2009) offered was that in the case of familiar sequences (such as numbers within a decade) people retrieve order information directly from long-term semantic memory.

¹ In the following, we will focus exclusively on symbolic comparison tasks (and will not discuss distance effects on the dot comparison task further). Besides methodological issues regarding the dot comparison task (e.g., Gilmore et al., 2013; Fuhs & McNeil, 2013; Szucs et al., 2013), another reason for this is that recent findings have questioned whether distance/ratio effects in the case of symbolic and non-symbolic comparison reflect similar processes (Lyons, Nuerk & Ansari, 2015).

In contrast with the above findings which suggest that comparison and order judgments rely on distinct processes, there is evidence that spatial representations are strongly implicated in both ordinal and comparison judgments. This raises the possibility that the representations that underlie ordinal and comparison judgments are similar or even shared, although they might be utilized in a different way depending on the type of question asked. Franklin et al. (2009) proposed that both numerical comparison and order judgments rely on an analogue, spatial mental number line representation (e.g., Restle, 1970). There is indeed empirical evidence for a link between performance on comparison tasks and performance on the number line task in which participants estimate the location of numbers on a spatial line (e.g., Booth & Siegler, 2006; Laski & Siegler, 2007; Lyons et al., 2014), as well as a link between numerical ordering skills and number line performance (Lyons et al., 2014).

Another piece of evidence for a link between spatialized number line representations and order processing skills comes from a training study. Kucian et al. (2011) trained children on a computer-based number line task, and they found that children's ability to judge whether three numbers are in ascending or descending order improved after the training. This again suggests a link between number line-type representations and ordering skills, although it should be noted that some of the number line trials involved arithmetic problems (i.e., additions and subtractions), whereas others involved assessing the numerosity of a set of dots. Thus, it is possible that the improvement in ordering skills resulted from the arithmetic training or training in dot enumeration, rather than from the number line training *per se*.

However, although these findings suggest that spatialized number line representations might be important for order processing (as well as number comparison), other claims have been made about how order information is stored or represented. For example, associative chaining models (e.g., Lewandowsky & Murdock, 1989), that build ordered lists from associations between pairs of items, offer an alternative. Recently, Caplan (2015) argued that

models that combine associative chaining and the coding of spatial position might be the most suited to explain performance on relative order tasks.

In summary, although some findings seem to hint at the possibility that comparison and order judgments utilize similar representations, the evidence so far is inconclusive. At least some dissociation can be expected on the basis of developmental (Lyons et al., 2014) and neuroscientific (Lyons & Beilock, 2013) evidence. Distance effects in the case of ordering and comparison judgments also show dissociations (Goffin & Ansari, 2016).

Are numerical comparison and order judgements based on domain-specific representations?

Beside the question of the independence of comparison and ordering skills, another important question is to what extent the processes that underlie numerical magnitude and order judgments are specific to the numerical domain. This question is particularly pressing because of the established links between arithmetic skills and numerical comparison and ordering judgments, raising the issue of the role of domain-general abilities in maths performance. As we described above, magnitude judgments, especially those involving symbolic numbers, have been widely linked to the representation of numbers in the form of a mental number line (i.e., a visuo-spatial representation of the relative position of numbers; for review, see de Hevia, Vallar, & Girelli, 2008). Canonical distance effects on symbolic number comparison tasks have often been taken to reflect this spatialized representation. However, a considerable amount of research suggests that distance effects are not specific to numbers. Distance effects have also been reported for letters of the alphabet (Jou & Aldridge, 1999). Moreover, Fairbanks (1969), Friedman (1984) as well as Seymour (1980) reported distance effects in the case of some types of judgments about the relations between months of the year, whereas Franklin et al. (2009) reported reverse distance effects in the case of ordinal

judgments about months when triads of months crossed the calendar year boundary. This latter finding mirrored the effect that they found in the case of number triads which crossed a decade boundary (see above).

Beside the distance effect, another well-known effect that has been associated with the existence of a mental number line is the SNARC effect² (i.e., Spatial Numerical Association of Response Codes). In their seminal study, Dehaene, Bossini and Giraux (1993) showed that when evaluating a number's parity status, shorter latencies were obtained for small numbers when participants responded with the left hand, whereas for large numbers latencies for right hand-responses were shorter. However, the SNARC effect has also been demonstrated in the case of other types of ordered items, such as months, letters (Gevers, Reynvoet & Fias, 2003), and even types of fruit listed in a specific order (van Dijck & Fias, 2011). Thus, if the SNARC effect is interpreted as reflecting the spatial coding of representations, such findings suggest similar spatial coding in other domains. In particular, distance and SNARC effects in non-numerical tasks involving months of the year are consistent with the abundant evidence that events in time are represented along a "mental time line" (see Bonato, Zorzi & Umiltà, 2012 for a review).

However, although distance and SNARC effects have been taken as evidence of linear spatial representation within a domain, there is in fact considerable debate over how such effects should be interpreted. With regard to distance effects, debate centres around whether such effects in comparison tasks reflect general aspects of decisional processes rather than the spatial nature of underlying representations (Lyons et al., 2015; Sasanguie, Defever, Van den Bussche, & Reynvoet, 2011; Van Opstal, Gevers, De Moor, & Verguts, 2008; Van Opstal & Verguts, 2011); with regard to SNARC effects, one key issue is whether such effects reflect visuo-spatial representations in semantic memory or temporary position-space associations in

² Although our study did not include an examination of the SNARC effect, we believe that the general issues that we raise in these sections are highly relevant to our investigation.

working memory (van Dijck & Fias, 2011; van Dijck, Gevers, & Fias, 2009). Thus, the fact that distance and SNARC effects are found in both numerical and non-numerical domains cannot straightforwardly be taken as evidence of common types of underlying representations. The unresolved issue of the similarity of numerical and non-numerical representations of sequences (and the processes that act on such representations) continues to attract considerable attention within both experimental psychology (Berteletti, Lucangeli, & Zorzi, 2012; Casarotti, Michielin, Zorzi, & Umiltà, 2007; Di Bono & Zorzi, 2013) and neuroscience (Attout, Fias, Salmon, & Majerus, 2014; Fias, Lammertyn, Caessens, & Orban, 2007; Zorzi, Di Bono, & Fias, 2011), with Berteletti et al. (2012) arguing that linear representations in the numerical domain become generalized developmentally into linear representations of other sequences such as months and letters (though see Hurst, Monahan, Heller, & Cordes, 2014).

The current study

Given this set of important but unresolved issues, the current study had the following, closely related, aims. First, we wished to investigate further the relationship between comparison and order processing performance in the case of numbers, and explore the relationship between these processes and arithmetic skills. Second, we hoped to obtain a better understanding of the nature of representations that underlie performance on number comparison and number ordering tasks. Finally, we wanted to explore the possibility that symbolic number comparison and number ordering skills are based on types of representations that are not specific to the domain of numbers. Based on earlier findings, the most likely candidate for such representations would be a mental line-type representation of the position of items.

1. **Exploring the link between number comparison and order processing, and arithmetic skills**

Regarding the first aim, we employed the following method. When exploring the relationship between comparison and order processing skills, we did not simply look at correlations between the tasks, but also the relationship between these tasks and arithmetic skills. Although recent studies have very usefully compared the predictive power of numerical comparison and order representations with regard to mathematical/arithmetic skills (Attout et al., 2014; Goffin & Ansari, 2016; Lyons et al., 2014; Vogel et al., 2014), the methodologies and participant samples of these studies have left several issues unresolved. Lyons et al. (2014) did not explore the relationship between numerical comparison and order processing skills. Moreover, their study was a developmental one conducted with children of different ages. Although this allowed the authors to observe important developmental changes in the predictive power of ordering versus comparison tasks, they did not include an adult sample, making it difficult to draw conclusions regarding mature mathematical skills. Similarly, Attout et al. (2014) were primarily concerned with the predictive power of comparison and order processing with respect to the development of maths abilities. Their longitudinal study was conducted with young children, and they found a progressive change in the processes underlying ordinal representations. Thus, the strongest conclusion that can be drawn from Attout et al.'s (2014) study is that number comparison skills mature earlier than order representations (see also Colomé & Noël, 2012, Lyons et al., 2014).

Goffin and Ansari's (2016) study involved adults, and it focussed on the distance effects observed in the case of comparison and ordering tasks. As mentioned above, these authors reported dissociations between the distance effects observed in the case of the two tasks, and at the same time, they established that distance effects in each of these tasks independently predicted variance in mathematical skills, after the effects of visual-spatial working memory and inhibition were controlled. Goffin and Ansari (2016) also argued that distance effects on these tasks were purer indicators of performance than scores derived from

accuracy and RTs, as these indices were less influenced by factors such as the speed of processing and differences in the task procedures.

Notably, this claim conflicts with those of Lyons et al. (2015) and Maloney, Risko, Preston, Ansari and Fugelsang (2010), who questioned the reliability of numerical distance effects. Lyons et al. (2015), in line with Verguts and van Opstal (2005), argued that distance effects might arise due to the relative frequency of numbers, and they simply indicate individuals' familiarity and fluency with which they can manipulate numbers. We return to this issue below.

In the current study, we used a symbolic number comparison task together with a number ordering task, and we also compared the predictive power of the two tasks with regard to arithmetic performance. We were interested in how performance on these tasks was related in adults. Based on earlier findings (Goffin & Ansari, 2016), we expected that performance on both the number comparison and the number ordering task should be strongly related to arithmetic performance in an adult group. We additionally tested the hypothesis that the effect of comparison skills on arithmetic abilities is mediated by number ordering skills. Given that ordinal processing becomes more important than comparison skills with development (Lyons et al., 2014), it is of interest whether comparison skills have any independent predictive value after their shared variance with ordering skills is taken into account.

2. Investigating the representations that underlie performance on the number comparison and number ordering tasks

In order to obtain a better understanding of the representations that underlie performance on the number comparison and number ordering tasks, we employed the following methods. First, we investigated the relationships between these tasks and performance on a number line task, with the latter task typically assumed to tap the

spatialized representation of the number series. In particular, we were interested in the possibility that numerical order processing skills might be supported by similar visuospatial representations as performance on the number comparison task. As we described above, number comparison (e.g., Booth & Siegler, 2006; Lyons et al., 2014) as well as number ordering skills (e.g., Kucian et al., 2011) have been linked to performance on the number line task. Although there is some controversy over how performance on the number line task should be interpreted (Hurst et al., 2014), such findings at least suggest that numerical order and comparison judgments might be related because they utilize a similar visuospatial representation of the relative position of numbers. So far only Lyons et al. (2014) have investigated numerical comparison, number ordering and number line performance together. They were concerned with the predictive power of performance on these tasks regarding mathematical abilities, and thus their analyses did not explore the relations between comparison/ordering and number line performance.

A second way that we investigated the processes and representations involved in numerical comparison and ordering performance was by investigating the presence of (both canonical and reverse) distance effects on these tasks. Whereas there is robust evidence for the presence of a canonical distance effect in the case of number comparison tasks, the findings regarding number ordering are somewhat mixed. Specifically, some studies (Franklin et al., 2009; Goffin & Ansari, 2016; Lyons & Beilock, 2013; Turconi et al., 2006) reported reverse distance effects in the case of tasks where participants had to make judgements about the order of items. However, Vogel et al. (2015) did not find this effect in young children, and Franklin et al. (2009) only found a reverse distance effect when number triads crossed a decade boundary. The current study carefully re-examined this issue.

3. Exploring the possibility that symbolic number comparison and number ordering performance are based on domain-general processes

A final issue that we wanted to address in the current study was whether the representations that underlie performance on the number comparison and number ordering tasks are specific to numbers. This question is of interest because Lyons et al. (2015) proposed that performance on basic numerical tasks (such as number comparison) might be linked to complex mathematical skills, because they measure individuals' familiarity and fluency with which they can manipulate symbolic numbers. To be able to investigate domain-specificity, beside the number ordering and number comparison tasks, we also administered ordering and comparison tasks that involved months of the year rather than numbers. Using these tasks, we addressed our research question in the following ways.

First, we investigated the link between month comparison and month ordering performance and arithmetic skills, and whether this link is mediated by number comparison or number ordering performance, respectively. If performance on month comparison and month ordering tasks are related to arithmetic skills in the same way as performance on the numerical equivalents of these tasks, then it may suggest that some type of domain-general skills underpin arithmetic.

Lyons and Beilock (2011) tried to address this question (at least for ordering skills) by including a letter ordering task as a control task for number ordering. These authors reported no relationship between letter ordering and arithmetic performance. However, we had some concerns over whether letter ordering is the best task to use to address the issue of domain specificity. Although the number sequence and the alphabet are, as Jou (2003, p.549) puts it 'the two most basic, overlearned and commonly used order sequences', unlike in the numerical system, ordinal information in the alphabet is not important in our most common everyday use of the system, with such information typically playing no functional role in how we use it (i.e., letters in the alphabet are used to generate words, for which ordinal position in

the alphabet is irrelevant). The conventional sequential order of the alphabet is only used in some exceptional cases (for example, when items are listed in alphabetical order).

By contrast, ordering in conventional systems marking time (days of the week, months of the year) is basic to the use of such systems. Moreover, there is considerable evidence that representations of such temporal systems are visuo-spatial in nature (Bonato et al., 2012; Friedman, 1983, 1984, 1989). Friedman (1986, 1992) argues that children initially learn the days of the week and months of the year by rote as verbal lists (just as they learn the alphabet), but that during adolescence they subsequently develop visuo-spatial representations of these systems as a result of using them in real-life problem solving tasks for which ordinal information is crucial (Friedman, 1986, 1992). That is, unlike in the case of the alphabet but similar to the case of numerical system, use of conventional time systems regularly recruits order processing skills. Friedman suggests that a visuo-spatial representation is typically recruited when making judgments about month order and position, but that adults retain and sometimes make use of verbal-list representations (Friedman, 1983, 1984). Given the robust evidence for the existence and use of visuo-spatial representations of conventional temporal systems, we decided to use a month comparison and ordering task that paralleled our number comparison and ordering tasks, and to examine whether performance on these tasks related to arithmetic ability in the same way as performance on their numerical equivalents.

A second way of investigating the domain-specificity or domain-generality of the representations that underlie performance on comparison and ordering tasks was to look at whether (either canonical or reverse) distance effects are present not only in numerical comparison and ordering but also in the case of the month comparison and month ordering tasks. Such similarities could provide an indication that comparison and ordering tasks rely on similar representations across domains. Franklin et al. (2009) reported reverse distance

effects in ordering tasks involving either numbers or months that crossed a decade/calendar boundary. However, as described above, other studies examining numerical ordering have produced somewhat different patterns of results (Lyons & Beilock, 2013; Turconi et al., 2006); moreover, these other studies did not include tasks involving months. Thus, it is potentially informative to re-examine the issue of the similarity of distance effects across these domains in both comparison and ordering tasks.

We are aware, though, of the need to interpret distance effects with caution. Given this, we employed one further method to investigate the representations that underlie participants' number/month comparison and number/month ordering performance: we investigated the relationship between their performance on these tasks and their thinking styles (i.e., a preference for verbal, spatial or object-visualizing thinking – see Blazhenkova & Kozhevnikov, 2009). Although this is a self-report measure, from which the findings also need to be interpreted cautiously, similar relationships between thinking style and performance on ordering/comparison tasks across the domains of numbers and time could strengthen the conclusions drawn from the results regarding the presence or absence of distance effects.

In summary, through investigating all of these questions together in a single study, we aimed to achieve a better understanding of the way numerical ordering and comparison performance are linked to arithmetic skills, the nature of the representations that underlie these skills, and whether these representations are specific to numbers or whether common representations underlie comparison and ordering performance across different domains.

Method

Participants

The participants were 87 undergraduate psychology students (69 females) aged between 18 and 56 years ($M = 21$ years 8 months, $SD = 6.12$). All participants provided written consent

and the study received ethical approval from the School of Psychology Ethics Committee. The students received ungraded course credit in return for their participation in the study. Due to the experimenter's error two participants did not complete the number comparison task, two participants did not complete the number ordering task, three participants did not complete the month ordering task, and another three participants did not complete the month comparison task. No participant missed more than 1 computer-based task. Additionally, one participant's data were excluded from the month ordering task, because their accuracy results showed that they responded randomly.

Materials

Tests of basic maths skills

Woodcock-Johnson III Tests of Achievement: Math Fluency subtest

The math fluency subtest of the Woodcock-Johnson III Tests of Achievement (Woodcock, McGrew, & Mather, 2001) was used as a measure of arithmetic skills. This test assesses the ability to solve simple addition, subtraction, and multiplication problems quickly. Participants were asked to work through a series of problems as quickly and accurately as possible within a three-minute time limit. The total number of correct items was calculated to provide a maths fluency score.

Number line estimation task

This task was based on the number-to-position problems used by Siegler and Opfer (2003) to examine numerical estimation abilities. In each problem, participants were presented with a number and asked to estimate where it would appear on a number line, where the position of '0' was marked on the left and '1000' was marked on the right. The task included 10 of these number line problems. The lines were positioned on the pages so that the starting points were not directly beneath one another, preventing the participants from using previous number lines to aid their estimations. Siegler and Opfer (2003) recommended that presented numbers

should be chosen so that their positions cannot be easily estimated based on prior knowledge.

The numbers presented were: 12, 48, 116, 263, 477, 526, 734, 805, 882, and 958.

Performance on this task was measured by calculating the total root-mean-square error. This is the square root of the average squared difference (in millimetres) from the position selected by the participant to the actual position of the number on the number line. Greater errors reflect poorer symbolic estimation performance. Cronbach's alpha for our sample was .54, which is an acceptable level of reliability.

Comparison tasks

A computerized number comparison task was administered. Participants were presented with a number between 1 and 9, and were asked to decide, as quickly as possible, whether it was smaller or larger than 5. They indicated a *smaller* or *larger* response using the keyboard.

A comparison task involving months (i.e., month comparison task) was also administered to participants. A month between January and September was displayed and participants were asked to decide whether it was before or after May in the calendar year. Eight practice trials were presented onscreen in both the number and month comparison tasks. These were followed by 48 trials, where each number between 1-4 and 6-9 was presented 6 times in a random order. The trials in the two tasks were matched so that for each number trial (e.g., 1, 2, 3), there was a corresponding month trial (e.g., January, February, March). Both accuracy and reaction times were recorded. Following Lyons et al. (2014) and Goffin and Ansari (2016) we used a composite of error rates (ER) and reaction times (correct trials only) to index performance on these tasks. The measures were combined according to the formula: $P = RT(1 + 2ER)$, where a higher value indicates worse performance. We preferred this method of creating composite scores over the method used by Lyons and Beilock (2011; i.e., using combined z scores of error rates and RTs), because our preliminary analyses suggested that accuracy and RTs were only weakly (negatively) correlated in the case of most of our tasks,

and they were uncorrelated in the case of the number comparison task. This formula also seems appropriate, because in the case of very high levels of accuracy the results basically reflect RTs only. Finally, Lyons et al. (2014) justified their choice of multiplying error rates by 2 in their formula on the basis that their tasks (just as ours in the current study) included binary forced choices (ER = .5 indicates chance).

Ordering tasks

A computer-based task was designed to measure numerical ordering ability (based on that of Lyons & Beilock, 2011). Participants were presented with number triads (e.g., 2, 4, 1), and they were asked to decide whether these triads were in increasing order, irrespective of the numerical distance between the numbers. All numbers were between 1 and 9. Participants indicated a *yes* or *no* response on the keyboard. A month ordering task was used to measure non-numerical ordering ability. In this task, participants were presented with the names of three months. They were instructed to decide whether these months were in the correct order with regard to their order within a calendar year. The first nine months of the year were included. Participants responded in the same manner as in the numerical ordering task. Eight practice trials were presented onscreen in both the number- and month-ordering tasks. These were followed by 48 trials. The trials in the two tasks were matched so that for each number ordering trial (e.g., 1, 4, 5), there was a corresponding month ordering trial (e.g., January, April, May). The trials were designed in a way so that the distance between the two extreme³ numbers/months within each triad was systematically manipulated. The smallest distance was 2 (i.e., when three consecutive numbers or months were included in the triad), and the largest distance was 7 (e.g., in the triads of 1, 3, 8 or February, April, September). Triads including the two extreme numbers and months (i.e., 1 together with 9, and January together with

³ Extreme items within each triad are the items that have the greatest distance between them when the items appear in the correct order (e.g., in the case of 2, 7, 4, the extreme items are 2 and 7, and the distance between them is 5). Thus, a distance of 2 means that the numbers were adjacent.

September) were not included, in order to avoid the possibility that participants respond to these trials using a response rule which does not necessitate the on-line checking of order within the triads. Additionally, to ensure that participants process all three items within the triads, in the case of incorrectly ordered triads the first two items always appeared in the correct order⁴. Both accuracy and reaction times were recorded. We used the same formula as for the comparison tasks to create composite scores from the accuracy and reaction time results.

Measure of thinking dispositions

Object-Spatial Imagery and Verbal Questionnaire

The 45-item Object-Spatial Imagery and Verbal Questionnaire (OSIVQ, Blazhenkova & Kozhevnikov, 2009) was administered to assess individual differences in cognitive style. Cognitive styles represent regularities in cognitive functioning, particularly in the acquisition and processing of information. The questionnaire is based on a three-dimensional model. Specifically, the questionnaire assessed whether participants had a preference for object imagery, spatial imagery or if they had a verbal cognitive style. People who have a preference for object imagery (referred to as object visualizers) use imagery to construct vibrant, clear images of individual objects. Blazhenkova and Kozhevnikov (2009) found that a dominance of this thinking style was characteristic of visual arts students. Additionally, students who scored high on this scale had a relatively low score on tasks measuring spatial abilities, such as the paper folding test and mental rotation. Individuals who rely mainly on spatial imagery (spatial visualizers) use imagery to represent and transform spatial relations. This thinking style is characteristic of science students, and it has been found to be associated with good performance on the paper folding and mental rotation tests (Blazhenkova & Kozhevnikov, 2009), as well as mechanical reasoning, and the ability to solve maths word problems

⁴ All number and month triads used in the study are listed in Table A in the Appendix.

(Morsanyi, Primi, Handley, Chiesi & Galli, 2011). Finally, those who possess a verbal cognitive style (verbalizers) mainly use verbal-analytical strategies when performing cognitive tasks. This thinking style is characteristic of humanities students, and it is associated with higher scores on vocabulary tests, and tests that require participants to formulate thoughts in a verbal format (Blazhenkova & Kozhevnikov, 2009). A key question that we wanted to address by the inclusion of this measure was the relationship between spatial and verbal representations and performance on the comparison and ordering tasks. The following is an example of an item from the object imagery scale: *‘When entering a familiar store to get a specific item, I can easily picture the exact location of the target item, the shelf it stands on, how it is arranged and the surrounding articles’*. An example of an item from the spatial imagery scale is: *‘I can easily imagine and mentally rotate three-dimensional geometric figures’*. Finally, the following item comes from the verbal scale: *‘When explaining something, I would rather give verbal explanations than make drawings or sketches’*. Participants rated their agreement with each item on a 5-point scale. The levels of agreement were ‘1 = Strongly Disagree’, ‘2 = Disagree’, ‘3 = Not Sure’, ‘4 = Agree’ and ‘5 = Strongly Agree’. Totals for each scale were calculated. Reliability analysis was also carried out for each subscale. In the current sample, Cronbach’s alpha for the object scale was .70, for the spatial scale it was .80, and for the verbal scale it was .76.

Measure of fluid intelligence

Raven’s Advanced Progressive Matrices (short form)

A short form of the Raven’s Advanced Progressive Matrices (Arthur & Day, 1994) was used as a measure of fluid intelligence. This test consisted of 12 items with three practice items from the Raven’s Standard Progressive Matrices (Raven, 1938) included at the start. This short form has been shown to have adequate psychometric properties, and it is a valid and reliable instrument (Chiesi, Ciancaleoni, Galli, Morsanyi, & Primi, 2012). In the present

sample Cronbach's alpha was .69. We included this measure because, potentially, performance on all of our measures could be related to participants' levels of fluid intelligence, which could explain some of the shared variance between our tasks.

Procedure

The participants completed all tasks in a single testing session which lasted approximately 50 minutes. They were tested in groups of 8-10. The order in which the tasks were presented was the same for all participants. The computer-based tasks were spread out so that they were presented at various intervals between the paper-and-pencil tasks. The participants first completed the timed math fluency test from the Woodcock-Johnson III Tests of Achievement. Then the month comparison task was presented onscreen. Following this, the OSIVQ was completed in the task booklet. Then participants were asked to complete the month ordering task on the computer. The number-line estimation task was then presented, followed by the number comparison task. Another booklet was distributed with the Raven's Advanced Progressive Matrices (short form) and participants were instructed to indicate their answers on the accompanying response sheet. This was followed by the final computer-based task: number ordering.

Results

First we present descriptive statistics for each measure (see Table 1). These show that on most measures there was a good variability in participants' scores. The exceptions were the comparison and ordering tasks, where accuracy rates were close to ceiling (especially on the comparison tasks). Nevertheless, there was a good variability in scores on the RT measures.

[Insert Table 1 here]

The relationship between performance on the comparison, ordering tasks and arithmetic skills

Our next analysis investigated the relationships between our measures of arithmetic skills, the comparison and ordering tasks, fluid intelligence and the measures of thinking dispositions (see Table 2). We report the correlations for the ordering tasks collapsed across correctly ordered and mixed-order trials, as our preliminary analyses suggested that performance across correctly ordered and mixed order trials within each task (i.e., number ordering and month ordering) were very highly correlated ($r_s > .82$), and the correlations between correctly ordered and mixed-order trials within each task and other measures were virtually the same. Nevertheless, later we will report some analyses where the presence of (reverse) distance effects on correctly ordered and mixed-order trials is investigated separately.

First we describe the first-order correlations between the (number and month) comparison and ordering tasks, and between these tasks and arithmetic skills. The correlations between the ordering and comparison tasks were very strong, ranging from .41 to .65. Performance on all of these tasks was also significantly related to arithmetic skills, as well as to fluid intelligence. Thus, in a number of analyses that we report below, we included fluid intelligence as a covariate. We conducted a regression analysis to confirm that both number comparison and number ordering explained unique variance in arithmetic skills. Indeed, when these tasks were entered into the regression equation together, both the number comparison ($\beta = -.26$ $p = .019$) and number ordering task ($\beta = -.41$ $p < .001$) were significant predictors of arithmetic abilities ($R^2 = .35$, $p < .001$)⁵. As the two performance measures were highly correlated, we included a multi-collinearity analysis for the regression analysis.

⁵ A stepwise regression analysis where the number and month comparison and ordering tasks were used as predictors of arithmetic performance yielded the same result, retaining the number ordering and number comparison tasks as significant predictors of arithmetic performance and excluding the month ordering and month comparison tasks.

According to the criteria proposed by Myers (1990), which specifies that a variance inflation factor (VIF) of ten or greater is cause for concern, the collinearity between comparison and ordering performance was at an acceptable level (VIF = 1.54).

[Insert Table 2 here]

Next, we conducted a mediation analysis to see if the effect of number comparison performance on arithmetic skills was mediated by number ordering ability. The aim of this analysis was to establish the extent to which the relationship between number comparison and arithmetic performance is explained by the shared variance between number ordering and number comparison performance. At the same time, this framework is also appropriate for establishing whether there is a direct (i.e., unmediated) link between number comparison and arithmetic skills once the shared variance with number ordering skills is accounted for⁶. The mediation hypothesis was tested using the INDIRECT procedure for bootstrapping (with 5000 bootstrap samples) to estimate 95% confidence intervals (CI; for more details see Preacher & Hayes, 2008). We estimated the indirect effect of number comparison on arithmetic skills, quantified as the product of the Ordinary Least Squares (OLS) regression coefficient estimating number ordering ability from number comparison (path *a* in Figure 1), and the OLS regression coefficient estimating arithmetic skills from number ordering controlling for number comparison (path *b* in Figure 1). A bias-corrected bootstrap-confidence interval (CI) of the product of these paths that does not include zero provides evidence of a significant indirect effect of number comparison on arithmetic skills through number ordering (Preacher & Hayes, 2008). The INDIRECT procedure revealed a significant indirect effect of number comparison on arithmetic skills through number ordering ($b = -.06$,

⁶ Using number comparison as the mediated variable is in line with developmental findings (Lyons et al., 2014). Whereas, in the case of young primary school children, number comparison performance is the best predictor of maths skills, in higher grade children, number comparison becomes the strongest numerical predictor of maths skills. Given that ordinal processing becomes more important than comparison skills with development, it is an important question whether, in the case of adult participants, comparison skills have any independent predictive value after their shared variance with ordering skills is taken into consideration.

95% $CI = -.010$ to $-.02$). Kappa squared was $.23$ (95% $CI = .10$, to $.34$) indicating a medium effect size for the indirect effect (Preacher & Kelly, 2011). Nevertheless, there was still a significant unmediated effect (c') of number comparison on arithmetic skills. That is, number comparison skills predicted arithmetic performance both directly (i.e., independent of number ordering skills) and indirectly (via number ordering skills).

Next we explored the question of whether the ordering skills implicated in arithmetic performance were specific to numbers (cf., Lyons & Beilock, 2011), by re-running the regression analysis and mediation analysis but replacing number ordering with month ordering. We found that, similar to findings of the regression involving number ordering, both the number comparison ($\beta = -.41$ $p < .001$) and the month ordering task ($\beta = -.23$ $p = .026$) explained unique variance in arithmetic skills ($R^2 = .28$, $p < .001$) when they were entered together into the regression equation. Collinearity between the predictor variables was at an acceptable level ($VIF = 1.20$). That is, ordering skills that were not number-specific predicted arithmetic ability independent of number comparison skills, confirming the relationship between arithmetic skills and non-numerical ordering skills. Note that this result is in contrast with the conclusions of Lyons and Beilock (2011) who claimed that the ordering skills implicated in mental arithmetic are specific to the domain of numbers. We also ran a mediation analysis to see if the effect of number comparison on arithmetic skills was mediated by month ordering. If this was the case, it would indicate that a general ordering ability that is implicated in number comparison performance explains some of the shared variance between number comparison and arithmetic skills. This question is theoretically important as, if this was the case, it would suggest that the link between number comparison skills and mental arithmetic can be attributed to non-numerical ordering abilities. The mediation hypothesis was tested using the INDIRECT procedure for bootstrapping (with 5000 bootstrap samples). The result was non-significant, indicating that the effect of number

comparison on arithmetic skills was not mediated by month ordering. Thus, although a general ordering ability (which is not specific to numbers) seems to be important for arithmetic skills, this general ordering ability does not explain the link between number comparison and arithmetic performance.

[Insert Figure 1 here]

An additional regression analysis was carried out to test the predictive value of number and month ordering skills with regard to arithmetic skills. When the two predictors were entered together, only number ordering skills ($\beta = -.47$ $p = .001$) were significant predictors of arithmetic abilities ($R^2 = .28$, $p < .001$). Collinearity between the two ordering tasks was at an acceptable level (VIF = 1.92). We followed up this analysis by a mediation analysis to test if the effect of month ordering skills on arithmetic ability was mediated by number ordering skills. This would suggest that the link between month ordering and arithmetic performance is present because of the demand for processing order information which is shared between number and month ordering. Additionally, this result would help in explaining the non-trivial fact that month ordering showed a relatively strong relationship with arithmetic skills. Once more we used the INDIRECT procedure for bootstrapping (with 5000 bootstrap samples) to test this hypothesis. There was a significant indirect effect of month ordering on arithmetic skills through number ordering (*unstandardized OLS regression coefficient* = $-.01$, 95% *CI* = $-.016$ to $-.004$), whereas the direct effect was not significant. Kappa squared was $.26$ (95% *CI* = $.08$, to $.42$) indicating a large effect size for the indirect effect (Preacher & Kelley, 2011). Including fluid intelligence as a covariate did not substantially change these results, and fluid intelligence was a non-significant covariate ($p = .29$). This result suggests that month ordering skills were linked to arithmetic abilities due to their shared variance with number ordering skills, and once this shared variance was accounted for, month ordering skills did not explain any additional variance in arithmetic abilities.

Finally, we also tested a regression model including number and month comparison as predictors of arithmetic skills. When the two predictors were entered together, only number comparison performance ($\beta = .40$ $p = .001$) was a significant predictor of arithmetic abilities ($R^2 = .24$, $p < .001$). Collinearity between the predictor variables was at an acceptable level (VIF = 1.43). We followed up this analysis by a mediation analysis to test if the effect of month comparison performance on arithmetic ability was mediated by number comparison performance. This would suggest that the link between month comparison and arithmetic performance is present because of a domain-general process which underlies performance on both the number and month comparison tasks. Again, such an analysis was important so that we could explain the non-trivial finding that month comparison skills showed a relatively strong relationship with arithmetic abilities. Once more, we used the INDIRECT procedure for bootstrapping (with 5000 bootstrap samples) to test this hypothesis. There was a significant indirect effect of month comparison on arithmetic skills through number comparison (*unstandardized OLS regression coefficient* = .03, 95% *CI* = .05 to .01), whereas the direct effect was not significant. Kappa squared was .20 (95% *CI* = .09, to .35) indicating a medium effect size for the indirect effect (Preacher & Kelley, 2011). Including fluid intelligence as a covariate did not substantially change these results, and fluid intelligence was a non-significant covariate ($p = .34$). This result suggests that month comparison skills were linked to arithmetic abilities due to their shared variance with number comparison skills, and once this shared variance was accounted for, month comparison skills did not explain any additional variance in arithmetic abilities.

Exploring the representations that underlie performance on the comparison and ordering tasks 1: Relationships with number line performance and thinking dispositions

One purpose of the following analyses was to investigate the relationship between comparison and ordering performance and performance on the number line task. All of the

comparison and ordering tasks were significantly related to number line performance, which hints at the possibility that performance on these tasks utilizes mental line-type representations. These correlations remained significant when we controlled for fluid intelligence (the partial correlation coefficients ranged from .24 to .50, all $ps < .042$).

Additionally, the correlational analyses indicated that both number and month comparison performance were related to spatial thinking styles. However, when we controlled for fluid intelligence, only the relationship between spatial thinking and the month comparison task remained significant. Moreover, object visualizing was negatively related to number ordering performance. These initial results showed some dissociation between the ordering and comparison tasks in terms of their relationships with the thinking style measures. Although verbal thinking styles were related to arithmetic skills, they were unrelated to performance on the comparison and ordering tasks.

In order to better understand the relationship between these variables, we performed a principal component analysis, which resulted in the extraction of three factors with eigenvalues over 1 (see Table 3). Together, these three factors explained 68% of the overall variance. Of these factors the most interesting one for our purposes was the first factor, which explained 40% of the variance, and which was dominated by the comparison and ordering tasks as well as by number line performance. Spatial thinking styles also loaded relatively strongly on this factor. Interestingly, number line performance also loaded strongly on a second factor, which was also characterized by low levels of spatial and verbal thinking styles. This suggests that although typically number line performance relies heavily on spatial (and to some extent verbal) strategies, some participants might have used alternative methods to perform this task. Finally, the last factor was dominated by the object visualizing thinking style. Regarding the processes that underlie (number and month) comparison and ordering

performance, the above results are in line with the notion that performance on these tasks utilizes visuo-spatial, mental line-type representations.

[Insert Table 3 here]

Exploring the representations that underlie performance on the comparison and ordering tasks 2: The presence of (reverse) distance effects in the comparison and ordering tasks

Our next analysis was aimed at establishing whether distance effects were present in the case of both the number and month comparison tasks, and if the effect of distance was similar across tasks (see Figure 2). A 2x4 ANOVA with distance from the reference number/month (5/May) and task (number/month comparison) as within-subjects factors indicated a significant effect of distance ($F(3,243) = 71.17, p < .001, \eta_p^2 = .47$), a significant effect of task ($F(1,81) = 84.57, p < .001, \eta_p^2 = .51$), and a significant interaction between distance and task ($F(3,243) = 6.84, p < .001, \eta_p^2 = .08$) on the composite scores. That is, performance was better for trials with a larger distance from the reference number/month, and it was also better for the number comparison task than for the month comparison task.

[Insert Figure 2 here]

In order to explore further the task by distance interaction we also conducted one-way ANOVAs with distance as a within-subjects factor separately for the two tasks. In the case of the number comparison task, there was a significant effect of distance ($F(3, 249) = 65.71, p < .001, \eta_p^2 = .44$). Additionally, post-hoc comparisons using Bonferroni-Holm corrections indicated that the difference between each particular distance was significant (i.e., all pairwise comparisons between different levels of distance showed a significant difference).

In the case of the month comparison task, a one-way ANOVA also indicated a significant effect of distance ($F(3,243) = 36.72, p < .001, \eta_p^2 = .31$). Post-hoc comparisons

using Bonferroni-Holm corrections indicated that the difference between each particular distance was significant (i.e., all pairwise comparisons between different levels of distance showed a significant difference), apart from the difference between the distances of 3 and 4 months, and 2 and 4 months. Thus, canonical distance effects were found for both number and month comparison tasks, although somewhat more consistently in the number task.

Our next analysis investigated the (reverse) distance effect in the case of the number- and month-ordering tasks. We analysed these effects separately for problems where the ordering was correct (see Figure 3), and where it was incorrect (see Figure 4), because previous studies (e.g., Lyons & Beilock, 2013) indicated that reverse distance effects were only present in the case of correctly ordered trials. First we analysed the results for correctly ordered trials. A 2x6 ANOVA on composite scores with distance between the two extreme items (2/3/4/5/6/7) and type of task (number ordering/month ordering) as within-subjects factors indicated a significant effect of distance ($F(5,400) = 5.38, p < .001, \eta_p^2 = .06$), and a significant effect of task ($F(1,80) = 388.13, p < .001, \eta_p^2 = .83$). That is, performance on the number ordering task was better than on the month ordering task, and trials with a larger distance between the two extreme items were more difficult than trials with a smaller difference between the two extreme items.

[Insert Figure 3 here]

However, there was also a significant interaction between distance and task ($F(5,400) = 4.52, p = .001, \eta_p^2 = .05$). Thus, we analysed the results separately for the two tasks using one-way ANOVAs with distance between the two extreme items (2/3/4/5/6) as a within-subjects factor. In the case of the month ordering task, there was a significant effect of distance ($F(5,400) = 5.88, p < .001, \eta_p^2 = .17$). Post-hoc comparisons using Bonferroni-Holm corrections indicated that performance was significantly better for the distance of 2 than for the distances of 5, 6, and 7. There were no other significant differences. In the case of the

number ordering task, the ANOVA indicated no significant effect of distance ($p = .175$). Overall these results provided some evidence for a reverse distance effect, but only in the case of the month ordering task.

We also ran the same analyses on the results for mixed-order trials. A 2x6 ANOVA with distance between the two extreme items (2/3/4/5/6/7) and type of task (number ordering/month ordering) as within-subjects factors indicated a significant effect of distance ($F(5,390) = 9.44, p < .001, \eta_p^2 = .11$), and a significant effect of task ($F(1,78) = 347.17, p < .001, \eta_p^2 = .82$) on the composite scores. The task by distance interaction was not significant ($p = .180$). That is, performance on the number ordering task was better than performance on the month ordering task, and trials with a larger distance between the two extreme items were less difficult than trials with a smaller difference between the two extreme items. Post-hoc comparisons using Bonferroni-Holm corrections indicated that performance was significantly worse for the distance of 2 than for the distances of 5, 6 and 7, and performance was significantly worse for the distance of 3 than for the distances of 4, 5, 6 and 7. There were no other significant differences. Overall these results showed that a canonical distance effect was present in the case of both the number ordering and month ordering tasks, and the effect of distance between items was similar across tasks.

[Insert Figure 4 here]

In summary, these analyses indicated that a (canonical or reverse) distance effect was present in the case of both correctly ordered and mixed-order trials. Nevertheless, the results were somewhat different for the month- and number-ordering tasks. In general, the month-ordering task was more difficult than the number-ordering task. In the case of correctly ordered trials, a reverse distance effect was found, but only in the case of the month-ordering task. In the case of mixed-order trials, a standard distance effect was present, and it was similar across the two tasks.

Discussion

In the discussion we will address each of our research questions separately, before we draw some general conclusions and discuss future directions for this line of research.

1. The link between number comparison and number ordering performance, and arithmetic skills

Our analyses regarding the link between comparison and ordering performance and arithmetic skills showed that the effect of number comparison and number ordering skills on arithmetic performance is overlapping, but only partially. We also revisited the claim of Lyons and Beilock (2011) that a specific numerical ordering ability is implicated in arithmetic performance, rather than domain-general ordering skills. Our results indicated that, in fact, arithmetic performance was also strongly linked to month ordering ability, and that month ordering skills explained unique variance in arithmetic skills when they were entered into the regression equation together with number comparison performance. However, our additional analyses showed that this non-numerical ordering ability did not mediate the effect of number comparison skills on arithmetic abilities. At the same time, the effect of month ordering ability on arithmetic skills was mediated by number ordering abilities.

Taken together, this pattern of results indicates that although number ordering and number comparison skills make independent contributions to arithmetic abilities, these tasks also utilize some shared underlying representations. On the basis of these findings, it is plausible to suppose that some shared basic underlying representations support both number comparison and ordering performance. Nevertheless, these representations are utilized in somewhat different ways in the case of comparison and ordinal judgments. We explore this issue further in the next section.

2. The nature of representations that underlie number comparison and ordering performance

As we described in the introduction, on the basis of previous research (e.g., Booth & Siegler, 2006; Franklin et al., 2009; Kucian et al., 2011; Lyons et al., 2014,) the most plausible candidate for a shared underlying representation for numerical comparison and order processing skills is a visuo-spatial mental number line. We explored this possibility by conducting a factor analysis including the comparison and ordering tasks, mental number line performance, and the thinking style measures. This analysis indicated that the comparison and ordering tasks loaded on the same factor as spatial thinking styles and mental number line performance, which is in line with this hypothesis. Although we recognize the limitations of the self-report measure of thinking styles, the fact that both spatial thinking style and number line performance loaded on the same factor suggest that this measure may capture something of interest. We also recognize that it is important to consider other explanations of why performance is correlated across tasks. In particular, because both the comparison and ordering tasks involve measurement of RTs, speed of processing might be an important common constraint on performance. However, the number line task, which loaded on the same factor in our analysis, did not involve RT measures, indicating that this factor does not just pick up speed of processing. Furthermore, the fact that these results generalized across the domains of numbers and months excludes the possibility that performance on the number comparison and ordering tasks is related just because of fluency of processing within the specific domain of numbers.

At the same time, as we mentioned in the introduction, previous research has also indicated that even if comparison and ordering performance do draw on the same visuo-spatial representation of the number line, this representation might be utilized in quite different ways in the case of comparison and order judgments. Some strong evidence for this claim is the presence of canonical distance effects in the case of comparison tasks and reverse distance effects in the case of order judgments (e.g., Franklin et al., 2009; Goffin & Ansari,

2016; Lyons & Beilock, 2013; Turconi et al., 2006). Nevertheless, as we also mentioned in the introduction, the evidence reported in previous studies regarding the presence of reverse distance effects in the case of ordering tasks is somewhat inconsistent. Thus, we revisited this question in the present study.

As expected on the basis of numerous earlier studies, the canonical numerical distance effect was present in the case of the number comparison task. Regarding the reverse distance effect, we found no evidence for this in the case of correctly ordered trials in the number ordering task. At the same time, on mixed-order trials a canonical distance effect was present. Thus, these results do not support the claim that the processes implicated in number comparison and number ordering judgments differ substantially, although the distance effect appears to be much weaker in the case of ordinal judgments. However, these results leave the possibility open that in the case of correctly ordered trials in the number ordering task both canonical and reverse distance effects might be present (cf. Franklin et al., 2009). That is, it may be that the lack of an observable distance effect in the case of such trials may be due to canonical and reverse distance effects cancelling each other out. Thus, the processes implicated in comparison and ordering performance might still be somewhat different.

The result that a reverse distance effect was not found in the case of correctly ordered number ordering trials is important, as there is a theoretical debate that centres on whether simple performance measures (such as accuracy and RTs) or measures derived from distance effects are the most appropriate for indexing performance on comparison and ordering tasks (e.g., Goffin & Ansari, 2016; Lyons et al., 2015; Maloney et al, 2010). In line with previous studies (Lyons et al., 2015; Maloney et al, 2010) that questioned the reliability of distance effects in the case of comparison tasks, our study has found no evidence for a reverse distance effect in the case of the number ordering task. We will return to these questions after discussing the distance effects in the case of the month comparison and ordering tasks.

3. Evidence regarding the domain-specificity vs. domain-generality of the processes that underlie performance in the comparison and ordering tasks

We will first consider the findings from the comparison tasks and then those from the ordering tasks. With regard to performance on the comparison tasks, there is converging evidence that the underlying processes are shared across domains. Indeed, the factor analysis indicated that the number and the month tasks both load on a factor that also includes number line judgments and spatial thinking styles (as well as performance on the ordering tasks). Additionally, the canonical distance effect was present in the case of both tasks. This is in line with earlier reports of the distance effect being present in the case of other tasks involving the months of the year (Fairbanks, 1969; Friedman, 1984; Seymour, 1980), and also with evidence that events in time are represented along a “mental time line” (see Bonato et al., 2012 for a review). An important implication of these findings is that if the distance effect is present in the case of months, then the distance effect in the case of numbers should not be interpreted as indicator of an approximate representation of quantities (i.e., as reflecting ANS functioning), although it is clearly linked to the relative positions of items. Indeed, in the case of the month comparison task, canonical distance effects were obtained although the task asked participants to judge the relative position of months (i.e., whether they were before or after May), rather than their magnitude. The conclusion that distance effects do not reflect ANS functioning is also supported by Lyons et al. (2015) who reported that individual differences in ratio effects (which are closely related to distance effects) on symbolic and non-symbolic comparison tasks were unrelated.

Whereas the relationship between performance on the month comparison task and visuo-spatial mental time line representations is not very surprising, the link between month comparison and mathematical skills requires some explanation. A plausible reason for this link could be that the ability to mentally “line up” items is necessary to create both mental

number lines and mental time lines, and mental line-representations support arithmetic abilities. Indeed, as we have pointed out, both numerical and month comparison skills were related to performance on the number line task. However, in the present study we did not find a link between number line performance and arithmetic skills. In fact, arithmetic performance was also not linked to spatial thinking styles. Thus, it seems that the link between comparison performance and arithmetic skills is based on some more specific requirements of the comparison tasks. Nevertheless, whatever this requirement might be, it is shared between the number and month comparison tasks, as indicated by the fact that the effect of month comparison performance on arithmetic skills was mediated by number comparison skills.

An alternative possibility is that performance on the comparison and ordering tasks, as well as the mental number line task is based on a shared underlying representation of ordered sequences in working memory (see van Dijck & Fias, 2011). Indeed, this interpretation is also compatible with our finding that spatial thinking styles are related to performance on these tasks, as spatial attention was found to be implicated even in serial order verbal working memory retrieval (van Dijck, Abrahamse, Majerus & Fias, 2013). Thus, when we discuss the link between comparison and ordering skills, and the ability to mentally line up items, we leave the possibility open that the ordered representations are only temporarily available in working memory.

A potential criticism regarding the interpretation of these results is that months are often represented using numbers, and, thus, it is possible that the shared variance between the number and month comparison tasks is simply explained by the fact that participants converted the months into numbers while performing the tasks. One reason to doubt that this is the correct explanation of the findings is that the difference between RTs for the number and month comparison tasks was very small (less than 120 ms on average – see Table 1), which makes it unlikely that participants had the time to convert the month names into

numbers. The instructions given in the month comparison task also did not ask for magnitude judgments; rather, participants were asked to judge whether a given month happened before or after May in the calendar year. Additionally, the distance and ordering tasks with numbers were completed by participants after they completed the same tasks with months, thus, they were not primed in the study to rely on numerical representations in these tasks. Finally, although the effect of month comparison skills on arithmetic performance was mediated by number comparison performance, the effect size was medium level, and the relationship between month comparison and arithmetic performance appeared to be weaker than the link between number comparison and arithmetic performance, which indicates that a substantial amount of variance was unshared across the number and month comparison tasks.

Regarding the ordering tasks, again we found evidence for a relationship with “mental line”-type representations and spatial thinking. Additionally, ordering skills as indexed by both number and month ordering performance were strongly linked to arithmetic performance. Moreover, the effect of month ordering skills on arithmetic performance was mediated by number ordering skills (and the effect size was large), which provides further evidence for the presence of shared underlying representations.

Nevertheless, the results regarding the presence of (reverse) distance effects on the number and month ordering tasks also point at some important differences between these tasks⁷. Specifically, in the case of the number ordering task, only a canonical distance effect was evident, and only in the case of mixed-order trials. By contrast, in the case of the month ordering task a reverse distance effect was present in the case of correctly ordered trials, whereas a canonical distance effect was found in the case of mixed-order trials.

A possible explanation for these patterns for the month ordering task (i.e., a canonical distance effect in the case of mixed-order trials, and a reverse distance effect in the case of

⁷This can be considered as additional evidence that months were not converted into numbers by participants.

correctly ordered trials) is that items that are in close proximity in a sequence (i.e., where the distance between items is small) trigger a “correct” response, perhaps due to strong associations between the names of the months. This would be in line with associative chaining models of ordered sequences (e.g., Lewandowsky & Murdock, 1989) where adjacent items are strongly connected by associative links. These associations would facilitate responding in the case of correctly ordered trials (which would explain the reverse distance effect), but would have to be inhibited in the case of incorrectly ordered trials. Although these effects could also be expected in the case of numbers, the association might be weaker, because numbers very often appear in non-canonical orders (for example, in the case of multi-digit numbers), whereas months more rarely appear in a non-canonical order. Indeed, it is an interesting possibility that, because maths typically involves dealing with numbers in non-canonical order, it might be important for the development of proficiency with numbers to be able to use numbers independently of their associations with other numbers within the count list. Because children start out learning numbers primarily within the context of the count list, it may be that associations between adjacent numbers are initially very strong. Such associations may be initially helpful in learning the single-digit number series, but are likely to be unhelpful when dealing with multi-digit numbers or performing numerical operations. Thus, it is probable that with experience of arithmetic and exposure to multi-digit numbers these associations are weakened. Two interesting predictions arise from these results. First, if this developmental account is along the right lines then reverse distance effects in ordering tasks should be present when the number ordering task is administered to young children, but the effect should gradually decrease with development. Another interesting possibility is that if ordering performance indeed requires inhibition skills, this could potentially explain why the relationship between performance on the dot comparison task and arithmetic abilities was mediated by number ordering ability in Lyons and Beilock’s

(2011) study (i.e., due to the shared inhibitory component of these tasks – see Fuhs & McNeil, 2013; Gilmore et al., 2013; Szucs et al., 2013). At the same time, this would also raise the possibility that it is not ordering skills *per se*, but inhibition skills that are necessary for arithmetic performance. Indeed, inhibition skills are very clearly implicated in arithmetic performance (e.g., Blair & Razza, 2007; Bull & Scerif, 2001; Swanson, 2011).

Although shared inhibition requirements could contribute to the link between ordering skills and arithmetic performance, it is unlikely that this link is fully explained by inhibition (see Goffin & Ansari, 2016). In fact, this link also does not seem to be explained by the fact that ordering abilities rely on visuo-spatial representations, as neither spatial skills nor number line performance correlated with arithmetic performance in the present study. Thus, the current results leave the question open what it is exactly about ordering skills that is important for arithmetic abilities.

A limitation of our study that deserves attention is that performance on the comparison and ordering tasks was, to a large extent, based on RT data. Thus, as already discussed, some of the shared variance between the comparison and ordering tasks could simply arise due to a common underlying relationship with processing speed. A good method to tackle this issue would have been to administer a separate measure of RT (as in Lyons et al., 2014). Nevertheless, this limitation of the design does not alter the conclusions that can be drawn regarding the relationships between these and other tasks. Indeed, performance on the comparison and ordering tasks was related to a task that was not based on RT measures (i.e., the number line task), and also to self-reported measures of thinking dispositions. More importantly, our mediation analyses showed differential patterns of shared variance among the comparison and ordering tasks, which makes it very unlikely that the relationship between the tasks simply arose due to the common RT component. Indeed, although all of these tasks were related to fluid intelligence (i.e., a task that taps a general processing requirement),

when the effect of fluid intelligence was controlled for, the results of the mediation analyses did not change.

Summary, conclusions and future directions

In the past decades, research has predominantly focussed on exploring the role of the ANS (e.g., Piazza et al., 2010) in symbolic number skills. These studies most often used symbolic and non-symbolic comparison tasks. Nevertheless, recent studies (starting with Lyons & Beilock, 2011 and Rubinsten & Sury, 2011) also highlighted the importance of number ordering skills in mathematical abilities. Indeed, Lyons et al. (2014) showed that beyond a certain point in development, number ordering abilities become the strongest predictors of arithmetic skills (although number comparison performance also continues to be a significant predictor). The current results substantially extend our previous knowledge regarding number ordering and number comparison skills. In particular, we provided new insight into the link between these skills, and also into the question of whether the link between mental arithmetic and ordering and comparison skills is specific to tasks that involve numbers.

Regarding the link between ordering and comparison abilities, our results indicate a partial overlap between these two skills. This is in clear contrast with Rubinsten and Sury (2011) who proposed that these skills represented two independent, parallel core systems. Our findings also prompt at least some reconsideration of claims about the independence of comparison and ordering skills (e.g., Lyons & Beilock, 2013; Goffin & Ansari, 2016). On the basis of the current results, we can also refute the claim that the ordering skills implicated in arithmetic performance are number-specific (Lyons & Beilock, 2011). In fact, our results additionally suggest that number comparison skills also rely on domain general processes, and are at least consistent with the suggestion that a spatial representation of item order underlies both ordering and comparison skills across domains (although such a representation

might be used in different ways in comparison and ordering tasks). Nevertheless, given that in the current study performance on the number line task was not significantly related to arithmetic performance, a link between comparison and ordering skills and visuo-spatial representations is not sufficient to explain why these skills are related to arithmetic skills (see also Goffin and Ansari, 2016 for similar conclusions). Indeed, our findings are more compatible with the proposal that performance on the comparison and ordering tasks, as well as the mental number line task is based on a shared underlying representation of ordered sequences in working memory (see van Dijck et al., 2013; van Dijck & Fias, 2011). Nevertheless, comparison and ordering skills might necessitate the ability to use ordered representations in a flexible way, depending on the instructions participants are given (i.e., to gather information about the order of items or the relative position of items along a continuum). Although the current findings extend our previous knowledge, a few questions remain unanswered. In line with Lyons et al. (2015), our results suggest that number comparison performance cannot be straightforwardly interpreted as a measure of magnitude processing skills (because there was a substantial association between performance on this task and the month comparison task, where the items do not represent magnitude information, although both tasks seem to rely on a spatial representation of item position). There is also a need to better understand the nature of the representations that underlie performance on the comparison and ordering tasks, and the involvement of working memory in performing these tasks.

Beside pointing at some important future directions for research, our findings can also inform the development of novel intervention methods. Indeed, theories of the component cognitive processes for mathematics are important because they can inform interventions to help improve mathematical abilities. Although some recently developed interventions for mathematics have focused on the ANS (see De Smedt, Noel, Gilmore & Ansari, 2013 for a

review) our findings suggest that it might be efficacious to use interventions that help support ordering abilities. Moreover, the current findings suggest that interventions that boost domain-general ordering abilities (rather than number ordering skills) might also offer significant benefits. Indeed, clinical observations suggest that individuals with dyscalculia struggle with following directions, and telling the sequence of past and future events. Children with dyscalculia have also been found to show impairments in transitive reasoning (Morsanyi, Devine, Nobes & Szucs, 2013), where they had to make judgements about the relative position of items. Interventions that use non-numerical materials to enhance ordering skills might be particularly useful for individuals who have pre-existing anxiety and negative feelings toward maths.

References

- Ansari, D. (2008). Effects of development and enculturation on number representation in the brain. *Nature Reviews Neuroscience*, 9, 278-291.
- Arthur, W. Jr., & Day, D. V. (1994). Development of a short form for the Raven advanced progressive matrices test. *Educational and Psychological Measurement*, 54, 394–403.
- 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:e92049.
- Attout, L. & Majerus, S. (2014). Working memory deficits in developmental dyscalculia: The importance of serial order, *Child Neuropsychology: A Journal on Normal and Abnormal Development in Childhood and Adolescence*, DOI: 10.1080/09297049.2014.922170
- Attout, L., Noël, M.-P., & Majerus, S. (2014). The relationship between working memory for serial order and numerical development: A longitudinal study. *Developmental Psychology*. 50, 1667-1679.
- Berteletti, I., Lucangeli, D., Zorzi, M., (2012). Representation of numerical and non-numerical order in children. *Cognition*, 124, 304–313,
- Blazhenkova, O., & Kozhevnikov, M. (2009). The new object-spatial-verbal cognitive style model: Theory and measurement. *Applied Cognitive Psychology*, 23, 638-663.
- 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.
- Bonato, M., Zorzi, M. & Umiltà, C. (2012). When time is space: evidence for a mental time line. *Neuroscience and Biobehavioral Reviews*, 36, 2257–2273.

- Booth, J. L., & Siegler, R. S. (2006). Developmental and individual differences in pure numerical estimation. *Developmental Psychology, 42*, 189–201.
- Caplan, J. B. (2015). Order-memory and association-memory. *Canadian Journal of Experimental Psychology, 69*, 221-232.
- Casarotti, M., Michielin, M., Zorzi, M., Umiltà, C., (2007). Temporal order judgment reveals how number magnitude affects visuospatial attention. *Cognition, 102*, 101–117.
- Chiesi, F., Ciancaleoni, M., Galli, S., Morsanyi, K. & Primi, C. (2012). Item Response Theory analysis and Differential Item Functioning across age, gender and country of a short form of the Advanced Progressive Matrices. *Learning and Individual Differences, 22*, 390-396.
- Colomé, À., & Noël, M. P. (2012). One first? Acquisition of the cardinal and ordinal uses of numbers in preschoolers. *Journal of Experimental Child Psychology, 113*, 233–247.
- Crawford, L. E., Regier, T., & Huttenlocher, J. (2000). Linguistic and nonlinguistic spatial categorization. *Cognition, 75*, 209–235.
- Dehaene, S., Bossini, S., & Giraux, P. (1993). The mental representation of parity and number magnitude. *Journal of Experimental Psychology: General, 122*, 371–396.
- 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.
- De Hevia, M.D., Vallar, G. & Girelli, L. (2008). Visualizing numbers in the mind's eye: The role of visuo-spatial processes in numerical abilities. *Neuroscience & Biobehavioral Reviews, 32*, 1361–1372.
- De Smedt, B., Noël, M.P., Gilmore, C., & Ansari, D. (2013). The relationship between symbolic and non-symbolic numerical magnitude processing and the typical and

- atypical development of mathematics: a review of evidence from brain and behavior. *Trends in Neuroscience and Education*, 2, 48-55.
- Desoete, A., Ceulemans, A., De Weerdt, F., & Pieters, S. (2012). Can we predict mathematical learning disabilities from symbolic and non-symbolic comparison tasks in kindergarten? Findings from a longitudinal study. *British Journal of Educational Psychology*, 82, 64–81.
- Di Bono, M.G. & Zorzi, M. (2013). The spatial representation of numerical and non-numerical ordered sequences: Insights from a random generation task *Quarterly Journal of Experimental Psychology*, 66, 2348-2362.
- Fairbanks, B. A. Jr. (1969). Experiments on the temporal aspects of number perception. *Dissertation Abstracts International*, 30, 403.
- Fias, W., Lammertyn, J., Caessens, B. & Orban, G. A. (2007). Processing of abstract ordinal knowledge in the horizontal segment of the intraparietal sulcus. *Journal of Neuroscience*, 27, 8952–8956.
- Franklin, M. S., Jonides, J., & Smith, E. E. (2009). Processing of order information for numbers and months. *Memory & Cognition*, 37, 644–654.
- Friedman, W. J. (1983). Image and verbal processes in reasoning about the months of the year. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 9, 650-666.
- Friedman, W. J. (1984). Analog and semantic models of judgments about the months of the year. *Memory and Cognition*, 12, 306-313.
- Friedman, W. J. (1986). The development of children's knowledge of temporal structure. *Child Development*, 57, 1386–1400.

- Friedman, W. J. (1989). The representation of temporal structure in children, adolescents and adults. In I. Levin & D. Zakay (Eds.), *Time and human cognition: A life-span perspective* (pp. 259–304). Amsterdam: North-Holland.
- Friedman, W. J. (1992). The development of children's representations of temporal structure. In F. Macar, V. Pouthas, & W. J. Friedman (Eds.), *Time, actions and cognition: Towards bridging the gap* (pp. 67–75). Dordrecht, The Netherlands: Kluwer.
- 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.
- Gevers, W., Reynvoet, B. & Fias, W. (2003). The mental representation of ordinal sequences is spatially organized. *Cognition* *87*, B87–B95.
- 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.
- 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.
- Halberda, J., Mazocco, M. M., & Feigenson, L. (2008). Individual differences in non-verbal number acuity correlate with maths achievement. *Nature*, *455*, 665–668.
- Holloway, I. D., & Ansari, D. (2009). Mapping numerical magnitudes onto symbols: the distance effect and children's mathematical competence. *Journal of Experimental Child Psychology*, *103*, 17-29.
- Hurst, M., Monahan, K. L., Heller, E., & Cordes, S. (2014). 123s and ABCs: Developmental shifts in logarithmic-to-linear responding reflect fluency with sequence values. *Developmental Science*, *17*, 892–904.

- Jou, J. (2003). Multiple number and letter comparison: Directionality and accessibility in numeric and alphabetic memories. *The American Journal of Psychology*, *116*, 543-579.
- 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.
- Kucian, K., Grond, U., Rotzer, S., Henzi, B., Schonmann, C., Plangger, F., Galli, M., Martin, E., & von Aster, M. (2011). Mental number line training in children with developmental dyscalculia. *NeuroImage*, *57*, 782–795.
- Laski, E. V., & Siegler, R. S. (2007). Is 27 a big number? Correlational and causal connections among numerical categorization, number line estimation, and numerical magnitude comparison. *Child Development*, *78*, 1723 – 1743.
- Lewandowsky, S., & Murdock Jr, B. B. (1989). Memory for serial order. *Psychological Review*, *96*, 25-57.
- Lyons, I. M., & Beilock, S. L. (2011). Numerical ordering ability mediates the relation between number-sense and arithmetic competence. *Cognition*, *121*, 256–261.
- Lyons, I.M. & Beilock, S.L. (2013). Ordinality and the nature of symbolic numbers. *Journal of Neuroscience*, *33*, 17052-61.
- Lyons, I. M., Nuerk, H. C., & Ansari, D. (2015). Rethinking the implications of numerical ratio effects for understanding the development of representational precision and numerical processing across formats. *Journal of Experimental Psychology: General*, *144*, 1021-1035.
- 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.

- Morsanyi, K., Devine, A., Nobes, A. & Szucs, D. (2013). The link between logic, mathematics and imagination. Evidence from children with developmental dyscalculia and mathematically gifted children. *Developmental Science*, *16*, 542-553.
- Morsanyi, K., Primi, C., Handley, S.J., Chiesi, F. & Galli, S. (2012). Are systemizing and autistic traits related to talent and interest in mathematics and engineering? Testing some of the central claims of the empathizing-systemizing theory. *British Journal of Psychology*, *103*, 472-496.
- Moyer, R. S., & Landauer, T. K. (1967). Time required for judgments of numerical inequality. *Nature*, *215*, 1519-1520.
- 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.
- Preacher, K.J. & Hayes, A.F. (2008). Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behavioral Research Methods*, *40*, 879–891.
- Preacher, K. J., & Kelley, K. (2011). Effect size measures for mediation models: quantitative strategies for communicating indirect effects. *Psychological Methods*, *16*, 93–115.
- Raven, J. C. (1938). *Progressive matrices: A perceptual test of intelligence*. London: H. K. Lewis.
- Restle, F. (1970). Speed of adding and comparing numbers. *Journal of Experimental Psychology*, *83*, 274–278.
- Rubinsten, O., & Sury, D. (2011). Processing ordinality and quantity: The case of developmental dyscalculia. *PLoS ONE*, *6*, e24079.

- Sasanguie, D., Defever, E., Van den Bussche, E., & Reynvoet, B. (2011). The reliability of and the relation between non-symbolic numerical distance effects. *Acta Psychologica, 136*, 73–80.
- Seymour, P. H. K. (1980). Internal representations of months: An experimental analysis of spatial forms. *Psychological Research, 42*, 255–273.
- Siegler, R. S., & Opfer, J. E. (2003). The development of numerical estimation: Evidence for multiple representations of numerical quantity. *Psychological Science, 14*, 237-243.
- 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.
- 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.
- Turconi, E., Campbell, J. I. D., & Seron, X. (2006). Numerical order and quantity processing in number comparison. *Cognition, 98*, 273–285.
- van Dijck, J.-P., Abrahamse, E. L., Majerus, S., & Fias, W. (2013). Spatial attention interacts with serial-order retrieval from verbal working memory. *Psychological Science, 24*, 1854–1859.
- van Dijck, J.P., Fias, W. (2011). A working memory account for spatial-numerical associations. *Cognition 119*, 114–119.
- van Dijck, J.P., Gevers, W., & Fias, W. (2009). Numbers are associated with different types of spatial information depending on the task. *Cognition, 113*, 248–253.
- Van Opstal, F., & Verguts, T. (2011). The origins of the numerical distance effect: the same–different task. *Journal of Cognitive Psychology, 23*, 112–20.

- Van Opstal, F., Gevers, W., De Moor, W., & Verguts, T. (2008). Dissecting the symbolic distance effect: Comparison and priming effects in numerical and nonnumerical orders. *Psychonomic Bulletin & Review*, *15*, 419–425.
- 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.
- Woodcock, R. W., McGrew, K. S., & Mather, N. (2001). *Woodcock-Johnson tests of achievement*. Itasca, IL: Riverside Publishing.
- Zorzi, M., Di Bono, M.G., Fias, W., (2011). Distinct representations of numerical and nonnumerical order in the human intraparietal sulcus revealed by multivariate pattern recognition. *NeuroImage*, *56*, 674–680.

Table 1. *Descriptive statistics for the measures of basic maths skills, the comparison and ordering tasks, the thinking style measures and the measure of fluid intelligence*

	Mean	SD	Minimum	Maximum
Arithmetic skills	110.15	21.24	67	158
Number line	3.57	1.68	1.21	9.61
Number comparison (acc.)	.98	.03	.75	1.00
Number comparison (RT.)	515.92	77.64	382.79	717.89
Month comparison (acc.)	.97	.03	.83	1.00
Months comparison (RT.)	631.20	132.70	444.51	1098.42
Number ordering (acc.)	.94	.05	.71	1.00
Number ordering (RT)	1082.96	276.76	599.85	1876.68
Month ordering (acc.)	.89	.09	.60	1.00
Month ordering (RT)	1896.55	449.43	1013.41	3159.48
Object-visualizing	52.18	9.46	36	90
Spatial thinking	41.99	8.08	23	65
Verbal thinking	46.40	7.89	21	63
Raven's matrices	5.06	2.30	0	11

Table 2. *Correlations between the measures of basic maths skills, the comparison and ordering tasks, fluid intelligence and the thinking styles measures (where it was necessary, scores had been converted so that higher scores on each task indicate better performance).*

	1.	2.	3.	4.	5.	6.	7.	8.	9.
1. Arithmetic skills	--								
2. Number line	.18	--							
3. Number comparison ¹	.50**	.51**	--						
4. Number ordering ¹	.53**	.40**	.60**	--					
5. Month comparison ¹	.39**	.25*	.55**	.56**	--				
6. Month ordering ¹	.42**	.32**	.41**	.69**	.64**	--			
7. Fluid intelligence	.04	.27*	.35**	.27*	.30**	.35**	--		
8. Object visualizing	-.13	-.21*	-.12	-.15	-.04	-.24*	.02	--	
9. Spatial thinking	.10	.17	.23*	.19	.31**	.08	.28**	-.17	--
10. Verbal thinking	.27*	-.12	.10	.17	.08	.11	-.03	-.05	.19

¹Performance on these tasks was indexed by a composite score created from accuracy and RTs (the formula is described in the Methods section).

Table 3. *Component matrix presenting the results of a factor analysis including the comparison and ordering tasks, number line performance and the thinking style measures (higher scores on each task indicate better performance or higher levels of the measured trait)*

	Component		
	1	2	3
Number comparison	.77	.09	-.07
Number ordering	.84	.05	-.17
Month comparison	.80	-.12	-.24
Month ordering	.78	.13	-.11
Number line	.61	.45	.20
Object visualizing	-.27	-.05	-.89
Spatial thinking	.42	-.56	.29
Verbal thinking	.20	-.80	-.03

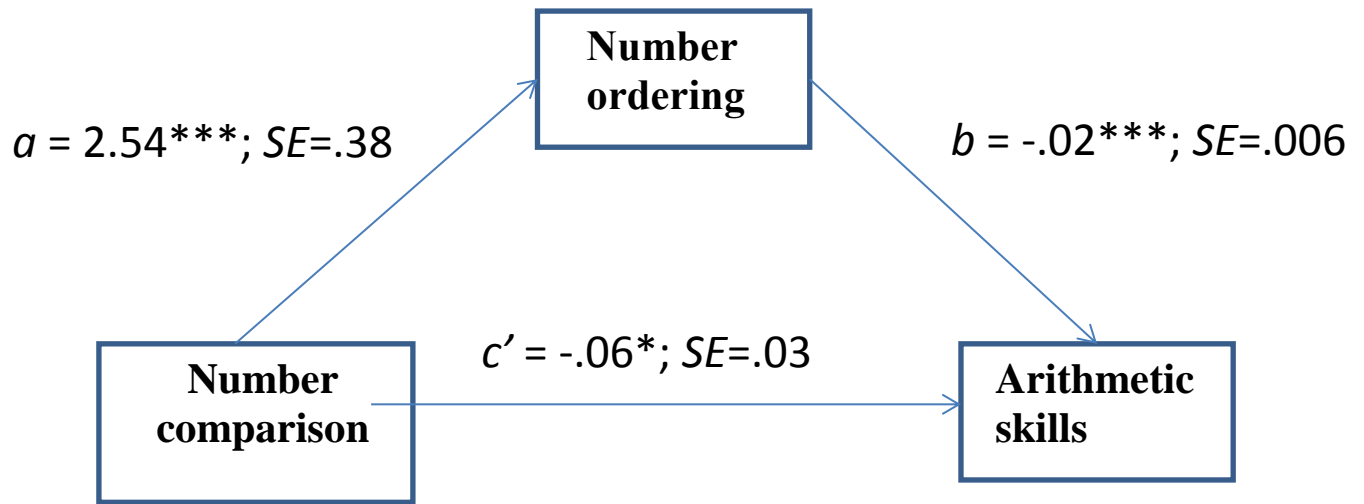
Figure captions

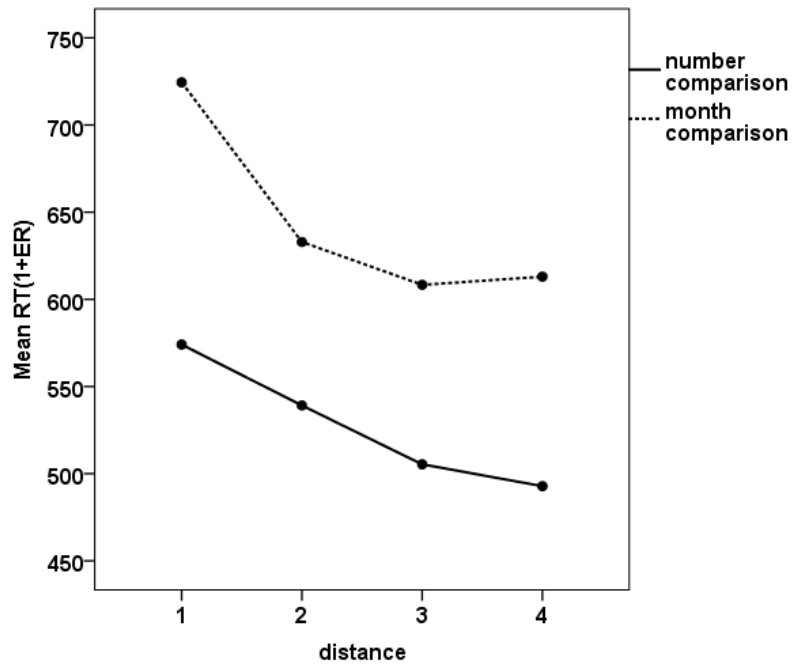
Figure 1. *Path model depicting the relationship between the number comparison task and maths fluency including number ordering skills as a mediator (a, b, c' are unstandardized OLS regression coefficients); * $p < .05$ ** $p < .01$; *** $p < .001$).*

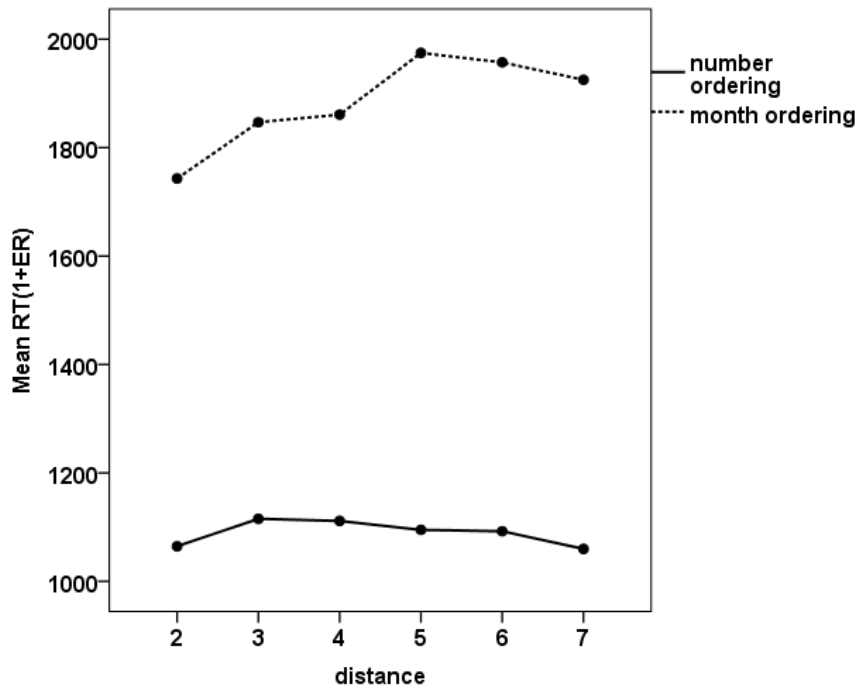
Figure 2. *Composite scores based on accuracy and RTs as a function of distance from the reference in the case of the number and month comparison tasks (higher scores indicate worse performance)*

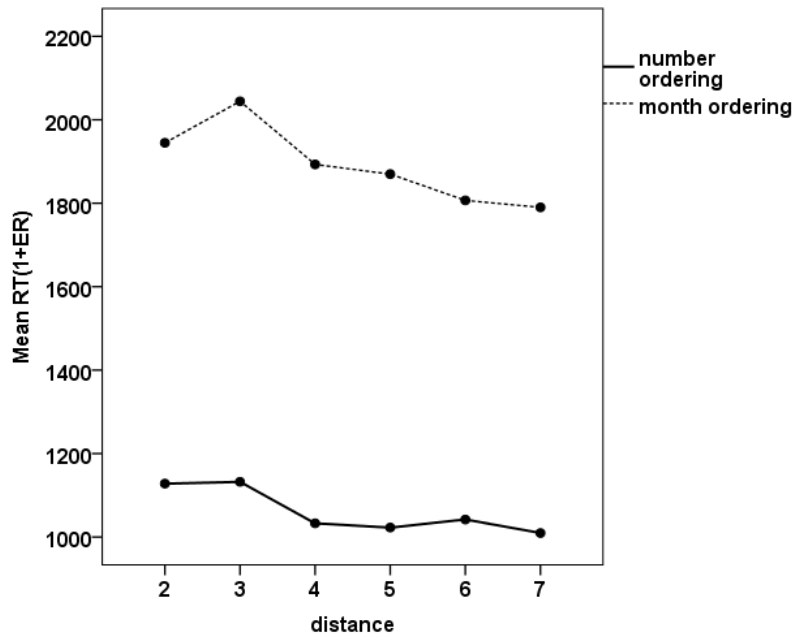
Figure 3. *Composite scores based on accuracy and RTs as a function of the distance between the two extreme items in the case of the correctly ordered trials on the number and month ordering tasks (higher scores indicate worse performance)*

Figure 4. *Composite scores based on accuracy and RTs as a function of the distance between the two extreme items in the case of the mixed order trials on the number and month ordering tasks (higher scores indicate worse performance)*









Appendix

Table A: Overview of the number and month triads used in the ordering tasks

Correct order		Mixed order		Distance
Numbers	Months	Numbers	Months	
<i>Practice trials</i>				
3 4 5	March April May	2 4 3	February April March	2
5 7 9	May June August	3 7 6	March July June	4
2 3 7	February March July	2 8 7	February August July	5/6
1 5 8	January May August	1 8 2	January August February	7
<i>Experimental trials</i>				
2 3 4	February March April	1 3 2	January March February	2
4 5 6	April May June	3 5 4	March May April	2
6 7 8	June July August	5 7 6	May July June	2
7 8 9	July August September	7 9 8	July September August	2
2 3 5	February March May	1 4 3	January April March	3
3 4 6	March April June	2 5 4	February May April	3
5 6 8	May June August	4 7 6	April July June	3
6 7 9	June July September	6 9 8	June September August	3
2 3 6	February March June	1 5 3	January May March	4
3 6 7	March June July	3 7 4	March July April	4
4 7 8	April May June	3 7 5	March July May	4
5 8 9	May August September	5 9 6	May September June	4
1 4 6	January April June	1 6 3	January June March	5
2 5 7	February May July	2 7 4	February July April	5
3 7 8	March July August	3 8 5	March August May	5
4 8 9	April August September	4 9 7	April September July	5
1 5 7	January May July	1 7 3	January July March	6
2 4 8	February April August	2 8 3	February August March	6
2 7 8	February July August	2 8 6	February August June	6
3 7 9	March July September	3 9 6	March September June	6
1 4 8	January April August	1 8 3	January August March	7
1 6 8	January June August	1 8 4	January August April	7
2 3 9	February March September	1 8 6	January August June	7
2 6 9	February June September	1 8 7	January August July	7