**RUNNING HEAD:** Processing and translating of numerical representations

Can adults automatically process and translate between numerical representations?

Iro Xenidou-Dervou a\*, Caroline Appleton a, Serena Rossi a, Natasha Guy a, Camilla Gilmorea

a Centre for Mathematical Cognition, Department of Mathematics Education, Loughborough University, Loughborough, United Kingdom

\*Corresponding author at:

Centre for Mathematical Cognition, Mathematics Education Centre, School of Science, Loughborough University*,* Loughborough, LE11 3TU, UK. T: +44 (0) 1509 223070, E: I.Xenidou-Dervou@lboro.ac.uk

**Contributions:**

CA, IXD and CG designed the study. CA, IXD and CG wrote the Stage 1 manuscript. IXD and CG revised the manuscript. NG and SR recruited participants and collected the data. SR and IXD performed the analyses. IXD, SR and CG wrote the Stage 2 manuscript. IXD, SR, CG and NG edited and reviewed the final manuscript.

**Conflict of interest statement:**

Authors declare no conflict of interest

**Data availability statement:**

Raw data, materials and analysis scripts are available on the Open Science Framework.

# Can adults automatically process and translate between numerical representations?

## Abstract

Arithmetic, and the ability to use numbers, is an important skill. Numbers can be represented in three ways: through number words, Arabic symbols or non-symbolically. Much research attention has focused on how associations form between these three numerical representations. However, it is not yet clear whether these associations are automatic or if they require Working Memory (WM) resources.

In this registered report, we used the dual-task paradigm to answer this question. Eighty-one adults were administered dot, digit, and cross-modal (i.e., dot vs. digit) magnitude comparison tasks in standalone and dual-task conditions with Phonological (PL) or Visuospatial (VSSP) WM interference. We found that all three types of magnitude comparison necessitated WM resources, albeit in distinct ways. Symbolic comparison necessitated VSSP WM. Surprisingly, in this task, accuracy improved under both WM interference conditions, evidencing the proposal that introducing executive function challenges in simple and familiar tasks can improve performance. In non-symbolic comparison, our findings demonstrated that the VSSP and the PL – albeit to a lesser extent – were employed. Finally, cross-modal comparison necessitated VSSP WM. These findings evidence the fundamental role visuospatial processing plays in numerical processing and that adults require WM resources for simply processing numerical representations and translating between them.

**Keywords**

Numerical cognition, Working memory, Numerical representations, Dual-task paradigm, Cross-modal comparison

Basic arithmetic skills, including counting and learning Arabic symbols, are acquired early in childhood. However, they have far-reaching consequences, including predicting future arithmetic skills, wider educational achievement (Duncan et al., 2007) and future socioeconomic status (Ritchie & Bates, 2013). Arithmetic is a skill that is important for everyday life, for example, for telling the time and buying food, and yet 25% of adults in the UK do not have the required numeracy skills for such day-to-day tasks (Department for Business, Innovation and Skills, 2011). In our daily lives, we encounter numbers in different forms; number words (the word “five”), symbolic (the Arabic symbol “5”), or non-symbolic (5 apples). Despite the importance of numbers and arithmetic skills in every aspect of life, it is not yet clear how we process numerical representations. The present study aimed to determine how far Working Memory (WM) and its components are involved in processing numerical representations.

## The nature of numerical representations

Much research attention has been focused on how we represent numbers and how the nature of these representations is related to arithmetic both in children and adults (e.g. Brankaer et al., 2014; Holloway & Ansari, 2009; Mundy & Gilmore, 2009). Here we use “representation” to mean internal representations, how numbers are represented cognitively and how these representations are linked together to provide meaning, rather than external physical representations of number. Numerical information can be represented in three ways: through words (often verbally), in a visual Arabic number form, or non-symbolically (Dehaene, 1992).

Number words are the first exact symbolic representation to be learnt in childhood. Children begin to recite the count sequence around their second birthday and by their third birthday begin to attach meaning to single-digit number words (Stock et al., 2009). Knowledge of the verbal count sequence is associated with success in later numeracy (Koponen et al., 2019) and is an important stepping stone to future arithmetic skills.

Like number words, Arabic symbols allow exact representations of quantities (Barner, 2018). They are a powerful tool, which allow us to concisely represent, access and manipulate exact numbers. Understanding of Arabic symbols is associated with arithmetic skills, both in children (e.g., Purpura et al., 2013; Vanbinst et al., 2018) and in adults (Orrantia et al., 2019). This stands to reason, as Arabic symbols are required to access the arithmetic curriculum in schools and to understand most numerical information presented to us. Furthermore, Arabic symbols encapsulate other mathematical constructs, such as place value, which are important for wider mathematical understanding (Moeller et al., 2011).

A third way that numbers can be represented is non-symbolically and research indicates there are two systems for this: one for small, exact numbers and one for large, approximate numbers. The small, exact system is known under various names across research, for example, the Object Tracking System (vanMarle et al., 2018) and is often associated with subitizing (Wender & Rothkegel, 2000). The subitizing range refers to the quantities which can be quickly and exactly enumerated, up to three in children and four in adults (Schleifer & Landerl, 2011). This system has been evidenced by research which finds that when enumerating a set of objects, accuracy decreases and reaction times increase significantly when the quantity increases above four (Revkin et al., 2008).

The system for processing large numbers has been referred to as analogue magnitudes or the Approximate Number System (ANS) (e.g., Dehaene, 1992; Feigenson et al., 2004; Halberda & Feigenson, 2008). The ANS is assumed to provide estimates of the quantity that a given set of non-symbolic stimuli represents. Repeated presentations of the same quantity result in slightly varying estimates, hence, mental representations of quantities are approximate in the ANS (Gallistel & Gelman, 2000). The precision of one’s ANS can be described as the reliability of activated estimates around the true quantity (Dietrich et al., 2015). In research, the ANS is commonly measured using dot comparison tasks; participants are presented with two dot arrays and asked to select which is the larger (e.g. Halberda et al., 2012). Correlational and experimental evidence in children and adults suggest that widely used ANS tasks require domain-general capacities, such as WM and inhibition skills (Gilmore et al., 2013; Guan, Gao, Li, Huang, & Si, 2021; Norris, et al., 2018; Xenidou-Dervou et al., 2014)

 It has been suggested that the precision with which individuals can represent and process non-symbolic quantities is associated with success in arithmetic. For example, Libertus et al. (2013) found that accuracy on a dot comparison task related to later arithmetic ability. However, the evidence for the relationship between the ANS and arithmetic is mixed (Schneider et al., 2017). Some research suggests that factors other than the numerosity of a set may influence the relationship with arithmetic (Gilmore et al., 2013), and other research suggests that there may be mediating factors, such as symbolic number knowledge (van Marle et al., 2014; Xenidou-Dervou et al., 2013).

Given that there are three representations of number, it raises the question of how these representations are connected, and whether it is in fact the connections between these representations that are more important for arithmetic, rather than the representations themselves. We now turn to a discussion of the research thus far into the connections between representations; here we present the evidence in the order that connections are thought to form in children (Jiménez Lira et al., 2017).

## Translating between numerical representations

Dehaene (1992) proposed the triple-code model as a way of explaining how numerical representations are related. The triple-code model describes the way numbers may be represented mentally in three different “codes”, what we refer to here as internal representations. As described above, numbers can be represented with number words (e.g., “three”), through Arabic symbols (3), or through an analogue magnitude code, which is a representation of quantity (Dehaene, 1992). These three representations of number can be linked together, allowing input in one representation and output in another. Throughout this study, we will use the phrase “translation” to describe the links between the different representations of a quantity. There is evidence from a range of sources that translating between numerical representations (i.e., intentionally converting or comparing quantities in different representations) is important for arithmetic abilities.

### Translating between number words and non-symbolic quantities

Translating between number words and quantity representations has been well-studied in small numbers. This association is often referred to as cardinality, the principle that non-symbolic quantities can be represented by symbolic quantities (Wynn, 1992). Young children are thought to gain this understanding around their fourth birthday (Batchelor et al., 2015; Gunderson et al., 2015). Hutchison et al. (2019) propose that because small quantities are processed exactly, for example through the Object Tracking System (Feigenson et al., 2004), they are processed more similarly to symbolic representations (either Arabic or number words) than to large quantities processed through the ANS. This may explain why forming associations between symbolic (Arabic or number words) and non-symbolic representations in small numbers is easier than in large numbers.

However, both children (Odic et al., 2015) and adults (O’Brien, 2014; Sullivan & Barner, 2013) struggle with translating non-symbolic quantities to number words in quantities outside the subitizing range. It is suggested that translating from a large approximate non-symbolic quantity (processed via the ANS) to an exact number word is cumbersome and may cause difficulties (Sullivan & Barner, 2013). The ability to form these associations is related to arithmetic (Odic et al., 2015) and therefore being able to represent these inaccurate non-symbolic quantities with a number word appears to be important. This highlights why we must consider the size of quantities (i.e. within or beyond the subitizing range) when examining the nature of numerical representations.

### Translating between Arabic symbols and number words

Translating between digits and number words has also been found to be related to arithmetic (Geary et al., 2000). Being able to provide a number word for an Arabic symbol was found to be related to later formal arithmetic achievement in kindergarten children (Purpura et al., 2013). Similarly, digit naming was the only factor which predicted growth in arithmetic in primary school children across a two-year period (Göbel et al., 2014), and these findings were replicated by Habermann et al. (2020). Number words and symbolic representations are both exact representations of number (Barner, 2018) and therefore accuracy in these tasks is often higher than in translations involving non-symbolic quantities.

### Translating between Arabic symbols and non-symbolic quantities

Less research has focused on translating between Arabic symbols and non-symbolic quantities; however, these associations are also related to arithmetic abilities. For example, Brankaer, Ghesquière and De Smedt (2014) found that children who were more accurate at matching dot arrays (non-symbolic quantities) to their Arabic symbols had higher arithmetic achievement.

These associations are often measured using cross-notation or cross-modal comparison tasks, where participants are presented with a symbol and a dot array and asked to select the larger. As with translating between non-symbolic quantities and number words, there is evidence that adults are particularly poor at these tasks (O’Brien, 2014). Izard and Dehaene (2008) found that when asking participants to estimate the numerosity of a dot array, participants significantly underestimated the true quantity. Furthermore, Lyons and colleagues (2012) found that reaction times were significantly higher when completing cross-notation tasks (translating between Arabic symbols and non-symbolic quantities) than when completing dot and digit comparisons (processing of non-symbolic quantities or Arabic symbols). These findings suggest that there may not be a direct association between Arabic symbols and non-symbolic quantities. In particular, we do not yet know whether adults can directly translate between Arabic symbols and non-symbolic quantities or whether access to number words is necessary for this process.

At present, research into translations between the different representations of number has not made the distinction between the two non-symbolic processing systems, the ANS and the OTS. For reasons highlighted above, primarily the differences between the small, exact system and the approximate large system, it is important to examine the nature of numerical representations in quantities of different sizes (i.e., within and beyond the subitizing range) separately.

## The nature of translations between number representations

The aforementioned literature establishes the importance of forming strong associations between the different forms of number representations for adults’ arithmetic skills, but little is known about the nature of these associations in adulthood. Several models attempt to explain the relationship between representations (Bernoussi & Khomsi, 1997) and particularly how representations come to gain meaning. In the triple-code model, as described above, semantic meaning of words and symbols is only provided through the connection with the non-symbolic quantity (Dehaene, 1992); this suggests that translations between representations are activated automatically to provide meaning.

The studies above focused on tasks where individuals intentionally translate or compare representations. Other research suggests that we may automatically translate one type of representation to another, even where it is not necessary for the task being undertaken. Studies have examined the automaticity of number processing using several methods. Reynvoet and Brysbaert (2004) used a priming study to investigate the automaticity of translations between Arabic and verbal representations. Participants were presented with either an Arabic digit or verbal number word (the prime) and then the alternative representation (the target) and asked to specify whether the target was odd or even. Where the prime and the target were numerically closer, response times for the parity judgement task were lower, suggesting that participants were automatically processing the numbers in their different modalities.

Automaticity of number processing has also been measured using congruency studies (Besner & Coltheart, 1979). In these studies, participants are asked to judge which is the physically larger of two Arabic digits, whilst ignoring numerical size. Where the numerically larger digit is also physically larger, the trials are congruent and reaction times are lower. However, where trials are incongruent, reaction times are higher (Reike & Schwarz, 2017). From this, it is inferred that participants are automatically accessing the non-symbolic quantity of the Arabic digit.

Furthermore, number words have been found to influence the processing of Arabic symbols in both adults and children, as seen in inversion effects, demonstrated in languages such as Dutch and German where number words are inverted (Xenidou-Dervou, Gilmore, et al., 2015; Zuber et al., 2009). This shows that representations of number in one modality can be influenced by a different modality, and that the processing of these representations may be automatic, i.e., that verbal representations are automatically activated when processing Arabic symbols, even where number words are not necessary (or relevant) to the task.

Neuroscientific studies have provided further evidence of the automaticity of letter and number processing. The processing of letters can be thought of in a similar way to number processing; both involve the association between a visual form (the letter shape or Arabic symbol) and a verbal sound (the letter or number sound). A neuroimaging study found that when congruent letters or numbers (i.e., the matching symbol and sound) were presented, patterns of brain activation were similar, and higher than when non-congruent pairs were presented (Holloway et al., 2015). Notably, the ability to form these automatic representations between letter-sound pairs has been found to relate to reading ability (Blau et al., 2010).

## The role of WM in numerical processing and translation

The studies described above considered automaticity in terms of the involvement of different numerical representations in tasks where they were not necessary. An alternative approach to automaticity is to consider the involvement of WM; where skills are automatised there is thought to be no WM involvement (Ding et al., 2017).

Working memory (WM) is a cognitive system where information is held and manipulated in the mind (Diamond, 2013). A commonly used theoretical model of WM is Baddeley and Hitch’s, a multi-component, limited-capacity system designed for storing and processing information (Baddeley & Hitch, 1974). It is thought to consist of the visuospatial sketchpad (VSSP) and the phonological loop (PL), which are responsible for storing information in the respective modalities (Baddeley, 2010), and the central executive (CE), which is responsible for processing information and regulates, controls and monitors the subsystems. It also contains the episodic buffer, which is responsible for combining information from the slave systems and from long-term memory (Baddeley, 2000). At present, the role of the episodic buffer in numerical cognition is not well understood and is not the focus of the present study.

 Correlational studies can provide indirect evidence about the role of WM in processing numbers in children (Friso-van den Bos et al., 2013; Xenidou-Dervou et al., 2013; 2014). Across multiple studies in school-aged children, the PL has been found to relate to symbolic abilities, including tasks such as counting, digit naming and symbolic comparison tasks (Östergren & Träff, 2013; Purpura & Ganley, 2014; Yang et al., 2020). Purpura and Ganley (2014) also found that the PL related to a non-symbolic comparison measure, whilst Yang et al. (2020) found the VSSP to be related to the ANS. These mixed findings provide evidence that WM is related to representing numbers and quantities. The above research is all in children. To the best of our knowledge, no correlational research has examined the role of WM in numerical processing in adults. Further, correlational studies cannot tell us whether WM resources are required for processing (i.e., comparing or manipulating numerical representations within a particular code: verbal, symbolic, non-symbolic) and translating (i.e., converting or comparing numerical representations across codes) numerical information, only that they are related.

Using the theory of WM, it is possible to use an experimental design to examine if certain components are *required* for the processing of numerical information. In studies using the dual-task paradigm participants complete a primary task (the task of interest, which is assumed to involve some aspect of WM), alongside a secondary, interference task known to involve a component of WM. If the primary task requires the component of WM being interfered with or suppressed by the secondary task, then performance on either the primary or secondary task will break down in comparison to control standalone conditions, i.e. without a dual-task load (for a review see Raghubar et al., 2010). Such an experimental design can evidence the causal role of WM in processing numerical information.

Few studies have so far used the dual-task paradigm to determine the role of WM in adults’ symbolic number processing (Herrera et al., 2008; Maloney et al., 2019; van Dijk, Gevers & Fias, 2009). In Maloney et al. (2019), adult participants completed a single-digit Arabic comparison task under two conditions, no load and phonological load. In the phonological load condition, participants were presented with a letter span before the comparison task, and then asked to recall the span after each comparison trial. Results showed that under the phonological load, performance in the symbolic comparison task was impaired in contrast to the no load condition, suggesting that the phonological loop is required in the processing of Arabic symbols. However, by only using a phonological secondary task, it is not possible to tell whether the effects found were due to the phonological interference specifically, or due to the increased cognitive load of completing two tasks simultaneously. Van Dijk, Gevers and Fias (2009) and Herrera et al. (2008) on the other hand, imposed both verbal and visuospatial WM load on adults’ symbolic magnitude comparison processing. In these studies, symbolic comparison was assessed with a task where participants see an Arabic digit ranging from 1 to 9 and must indicate whether the number they saw is smaller or larger than 5. Performance in this type of task elicits the so-called Spatial Numerical Association of Codes (SNARC-effect; Dehaene, Bossini, & Giraux, 1993), which reflects an association between numerical magnitude and response side, such that larger numbers are associated with the right side and smaller with the left. In both studies, under the spatial – but not the verbal – load the expected SNARC effect was not observed. These findings demonstrate that the VSSP may play a role when processing the spatial representation of number. Given the key differences in the primary task used across these studies, the question remains: Which component of WM is necessary when processing and translating between different number representations?

## The present study

The present study investigated the processing of Arabic symbols and non-symbolic quantities, and the role of verbal representations in translating between symbolic and non-symbolic representations. Using a robust, dual-task design we can determine which WM components are involved in the processing and translation of numerical representations. If associations between representations are processed automatically, then we would expect to see no WM involvement.

 To examine the processing of numerical representations, we administered dot comparison, digit comparison and cross-modal comparison tasks as primary tasks, which were conducted in standalone and dual-task (phonological and visuospatial) conditions. This allowed us to compare performance under PL and VSSP interference, ensuring that any detriment observed in task performance is due to the targeted WM component interference.

Examining performance across all three primary tasks allowed us to draw conclusions about the specific nature of numerical representations both when processing and translating different representations. The use of three comparison tasks allowed us to draw conclusions about the nature of each representation and ensured that any WM involvement would be due to the specific representation and not to the act of comparing any two quantities. If performance on the cross-modal task is impacted by the dual-task conditions but performance on the digit comparison and dot comparison tasks are not, then we know that WM is required for the process of translation, and not for simply processing the numerical representations themselves.

This method allowed us to answer further questions about the nature of representations in each modality. We expected to see phonological involvement in the symbolic comparison task, however, previous research is less clear about the WM involvement in dot comparison tasks, and therefore we aimed to clarify this finding. Maloney and colleagues (2019) found phonological involvement in a cross-modal mapping task, however they did not investigate VSSP involvement.

As discussed, non-symbolic numbers are processed through two different systems, the ANS for large numbers and the small exact system for small numbers. Therefore, to fully understand the translation of non-symbolic quantities to number words and Arabic representations, we must consider both non-symbolic representational systems. The present study therefore also examined the differences in how small (1-4) and larger (5-9) quantities are processed and translated. We chose these quantities, rather than quantities greater than 10, as whilst the non-symbolic representations are approximate, it is still possible for adults to attach Arabic symbols to these quantities. We expected that quantities in the small range would involve more phonological processes than those in the large range because small non-symbolic representations are assumed to be processed in a similar way to symbolic representations (Hutchison et al., 2019).

To address our primary research question, we designed secondary tasks that could interfere with the PL or VSSP components of WM. We aimed to address the following research questions:

1. Are symbols and non-symbolic representations accessed automatically or does access require the involvement of WM components?
	1. We hypothesised that the processing of Arabic digits will require the involvement of the phonological loop.
	2. We hypothesised processing of non-symbolic quantities will require the involvement of the VSSP
2. Can adults translate between symbolic and non-symbolic representations automatically or does translation require access to verbal representations?
	1. We hypothesised that translation between symbolic and non-symbolic representations will require access to the phonological loop.
3. Does the processing of numerical information differ for small and large quantities?
	1. We hypothesised that for symbolic processing and cross-modal translation, there would be no differences between small and large quantities and that both would require access to the phonological loop.
	2. We hypothesised that for non-symbolic quantities, small quantities will be processed automatically and large quantities will be processed using the VSSP.

The Stage 1 manuscript (<https://osf.io/z9cv2>) and related resources (e.g., stimuli, experiment scripts) for the following experiment can be found here: https://osf.io/ktq8e/

## Method

### Participants

Eighty-one adult participants (*Mage* = 25.32; *SD* = 11.09) took part in the experiment (n = 50 identified as women, n = 30 as men, and n = 1 responded “other”). Adult participants (age 18-65) were recruited via university email and social media (Nstudents = 63, Nother = 18). Research has shown that there is relatively little change in adults’ WM performance within this age-range (Alloway & Alloway, 2013).Participants had normal or corrected-to-normal vision and hearing and spoke English as their first language. Ethical approval was granted by Loughborough University Ethics Committee and participants were reimbursed for their time.

### Power Analyses

We conducted a-priori power analyses to calculate the required sample size using G\*Power (Faul et al., 2009; Lakens et al., 2022). Prevailing theories of number processing such as the ANS and Triple Code model have been developed for explaining individual differences in *accuracy*. Therefore, we based our power analysis on our assumptions of the minimum effect size of interest in accuracy (Lakens et al., 2022).

All power analyses were calculated using an alpha level of 0.05 and a minimum power of 90%. For primary tasks, there were a total of 160 trials. We calculated our minimum effect size of interest by considering what we believe to be the smallest relevant decrease in performance. Previous studies have demonstrated that adults’ accuracy rate on standalone comparison tasks of this type is very high (e.g., dot comparison accuracy: 99.7%, *SD* = 0.3 in Lyons et al. 2012). Given the expected high performance, we decided to power our study so that we could detect a difference of 5 out of 160 trials on the primary task; this would reflect a 3% difference, which we believe would be a meaningful decrease in performance. Based on these calculations the largest required sample size was N = 81, and therefore this was the sample size that was recruited for this experiment. For RT, this would allow us to detect differences of 50ms for the symbolic and non-symbolic comparison conditions and a difference of 80ms for the cross-modal comparison condition. For secondary tasks, there was a total of 20 trials. A decrease in 1 sequence length (e.g., remembering 6 items versus 5 items) would mean a decrease of 4 out of the 20 trials of the secondary task. We expect adults to remember on average up to 6 items (Monaco et al., 2013), therefore we expected the mean for the PL secondary task to be 16 (the number of correct sequences out of 20 if one correctly recalls 6 items). We decided that a drop of 4 out of 20 trials (i.e., 1 sequence length) would be a meaningful decrease in performance for our experimental design. Although calculations of our smallest effect sizes of interest were informed by our theoretical predictions and practical considerations, they could also inevitably be considered arbitrary since no study has previously examined these effects, therefore nonsignificant results should be treated tentatively. Calculations of effect sizes can be found in Appendix A, outputs for the largest power analysis can be found in Appendix B, and all other outputs can be found in the dedicated Open Science Framework (OSF) directory: <https://bit.ly/3lFeWll>.

### Materials

#### Primary tasks.

***Numerical comparison tasks****.* Participants completed symbolic, non-symbolic and cross-modal comparison tasks. The quantities used in each task were the same. Small numbers comprised 1-4, and large numbers 5, 7 and 9. These numbers were selected to ensure that the ratios between the numbers were large enough for participants to make judgements about which is larger using non-symbolic representations, and to equate the ratios across the small and large numbers. All unique combinations of these number pairs within sizes (small exact vs ANS) were used, with the exception of pairs with a ratio of 0.25, which were removed in order to equate difficulty across the small and large sets. 11 was added to the large set, to ensure that participants do not always select 9 as the larger quantity, however these trials were excluded from analysis. Further details about the quantities can be found on OSF. In the cross-modal comparison task, the side of presentation for the Arabic symbol was counterbalanced.

Quantities were presented on the screen, and participants were instructed to select the larger quantity and respond using the keyboard (“z” if the left quantity was larger, “m” if the right quantity was larger). Quantities appeared on the screen for 1000ms, to prevent counting, however participants could respond indefinitely. Dot arrays were created using MatLab and we controlled for visual properties such as surface area. Comparison pairs were created such that across all trials, no one property of the arrays (diameter, surface area, convex hull, density or contour length) would allow 100% accuracy. In half of the trials, visual parameters were congruent with quantity (i.e. the array with the higher quantity of dots also had larger diameter, greater density etc.), and in half of the trials, visual parameters were incongruent with dot quantity (Wang et al., 2020).

#### Secondary tasks.

***Phonological.***A reverse letter span task was used as a secondary task to load the PL component of WM. The sequence of events was as follows: 1) Participants were presented with a randomised sequence of letters (1 second per letter, presented orally through the computer) and told to remember the sequence. Each letter could only appear once in a sequence. Letters were chosen from the set “F, H, J, K, L, N, P, Q, R, S, T, Y”, as used in Maloney et al. (2019), 2) After completing eight trials of the primary task (approximately 8 seconds), participants were then asked to recall the sequence in reverse, with the response being entered into the computer by the experimenter. By recalling the sequence in reverse, it required participants to use their WM to process the information, as opposed to simply maintaining the letters in short-term memory.

The span ranged from three to seven letters, increasing in length throughout each condition, as this was found to be the range that an average adult can remember in a standalone reverse span task (Monaco et al., 2013). Four trials were used for each span length, resulting in a total of 20 trials. For the secondary task, we recorded accuracy of recall.

***Visuospatial.***A visuospatial span task was used as a secondary task to load the VSSP, which is an adapted version of a Corsi blocks task (Kessels et al., 2000). Participants were shown nine blue squares on the computer screen (see Figure 1). The blocks then changed colour individually (changing red for 1 second, then reverting to blue), which indicated a sequence (see video on OSF - https://bit.ly/3lFeWll). The blocks remained in the same positions on the screen for the length of the experiment. As in the verbal secondary task, sequence length ranged from three to seven items and increased throughout each condition of the primary task, with four trials for each span length.

After completing the primary task, the blocks were presented again. It was important that participants responded in the same manner in both the PL and VSSP dual-task condition, so that they are comparable. It was also important that the response mode for the primary and secondary tasks were different, to ensure that we are isolating the processing mode rather than the response mode. Participants responded to the primary task with their hands and therefore, participants responded verbally to the secondary tasks.

B

I

F

G

A

D

H

C

E

Figure 1. *VSSP secondary task. Above image shows the block presentation at the start of each trial. Below image shows the blocks at the end of the trial, with letters added to allow the participant to recall the sequence verbally.*

To allow participants to respond verbally, each square was labelled with a letter and the participant indicated to the experimenter the order of the sequence in reverse (see Figure 1). The location of the letters was randomly generated for each trial. This prevented participants from using their PL to rehearse the visual sequence whilst completing the primary task because the phonological response mechanism was only involved during recall. Again, for the secondary task, we recorded accuracy of recall.

### Procedure

All participants completed all conditions across two sessions. The order of primary tasks was counterbalanced across participants. The order of secondary tasks was also counterbalanced across participants but remained constant for individual participants across the two sessions[[1]](#footnote-2)*.* This means that participants completed each primary task in standalone and dual-task conditions, before moving on to the next primary task. Participants also completed both secondary tasks as standalone. An example of the procedure demonstrated for the non-symbolic comparison condition is shown in Figure 2.



Figure 2. *Example of conditions for non-symbolic comparison. Condition (a) shows standalone primary, (b) shows phonological dual-task and (c) shows visuospatial dual-task.*

### Data Analysis

 A factor to note when analysing dual-task performance is the trade-off between primary and secondary performance. Therefore, for each research question, we examined performance in both tasks, in comparison to the standalone conditions.

## First, for each participant, we calculated the mean accuracy and the Median RT (for correct trials only) for the primary task and the accuracy in the secondary task. Before conducting our main analysis, we performed normality checks. We plotted the data and examined skewness and kurtosis values. Following recommendations (e.g., Kline et al., 2011), we conducted non-parametric paired comparisons (Wilcoxon signed-rank) instead of parametric paired t-tests if skew was > |3| or kurtosis was > |4|. Outliers were examined for performance on each task (i.e., primary and secondary tasks) and extreme outliers (> 3.29 *SD*, Field, 2016) were removed from the respective analyses. All analyses were conducted using R Statistical Software (v4.2.2; R Core Team 2021).

Research questions 1 and 2 were answered via a series of planned paired comparisons to look for a) differences in primary task performance between standalone and dual-task conditions and b) differences in secondary task performance between standalone and dual-task (interference) conditions. These were performed separately for the different primary tasks (non-symbolic comparison, symbolic comparison, cross-modal comparison) and secondary tasks (PL, VSSP). Analyses were conducted for all trials combined before further analyses considered small and large trials separately, to answer Research Question 3. A summary of the preregistered analysis plan can be found in Appendix A, the detailed analysis in Appendix C and the study’s Stage 1 registered report on OSF (<https://osf.io/z9cv2> ).

## Results

After the removal of extreme outliers for each individual variable, we inspected the descriptive statistics for accuracy and median RT on all the primary tasks (Symbolic, Non-symbolic and Cross-modal) for all trials combined (see Table 1) and small and large quantities separately (see Table 2) for the standalone and each interference (PL or VSSP) condition.

**Table 1.**

*Descriptive statistics for primary and secondary task performance (accuracy and RT in seconds) in all experimental conditions.*

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |  |
|  |  | **N** | **Mean (SD)** | **Min** | **Max** | **Skewness** | **Kurtosis** |
| **Non-Symbolic primary task standalone** | Accuracy | 80 | 0.94 (0.04) | 0.78 | 1 | -1.48 | 2.77 |
| Median RT | 79 | 0.53 (0.09) | 0.37 | 0.81 | 0.64 | 0.18 |
| **Non-Symbolic primary task with PL secondary task** | Accuracy | 80 | 0.94 (0.04) | 0.83 | 0.99 | -1.24 | 2.99 |
| Median RT | 79 | 0.55 (0.09) | 0.38 | 0.82 | 0.80 | 0.31 |
| **Non-Symbolic primary task with VSSP secondary task** | Accuracy | 80 | 0.93 (0.04) | 0.77 | 0.99 | -1.45 | 2.98 |
| Median RT | 79 | 0.54 (0.09) | 0.38 | 0.82 | 1.03 | 1.46 |
| **Symbolic primary task standalone** | Accuracy | 79 | 0.97 (0.03) | 0.87 | 1 | -1.29 | 2.31 |
| Median RT | 79 | 0.47 (0.08) | 0.37 | 0.75 | 0.91 | 0.86 |
| **Symbolic primary task with PL secondary task** | Accuracy | 81 | 0.97 (0.02) | 0.89 | 1 | -1.39 | 1.82 |
| Median RT | 79 | 0.50 (0.07) | 0.37 | 0.69 | 0.49 | -0.11 |
| **Symbolic primary task with VSSP secondary task** | Accuracy | 81 | 0.97 (0.03) | 0.86 | 1 | -2.14 | 5.92 |
| Median RT | 79 | 0.50 (0.07) | 0.38 | 0.68 | 0.49 | -0.07 |
| **Cross-modal primary task standalone** | Accuracy | 80 | 0.91 (0.05) | 0.76 | 1 | -0.85 | 0.61 |
| Median RT | 79 | 0.63 (0.12) | 0.4 | 1.16 | 1.81 | 5.85 |
| **Cross-modal primary task with PL secondary task** | Accuracy | 79 | 0.90 (0.07) | 0.66 | 1 | -1.32 | 3.11 |
| Median RT | 79 | 0.63 (0.12) | 0.42 | 1.17 | 1.74 | 5.11 |
| **Cross-modal primary task with VSSP secondary task** | Accuracy | 79 | 0.90 (0.05) | 0.70 | 1 | -0.87 | 1.40 |
| Median RT | 79 | 0.63 (0.10) | 0.4 | 0.92 | 0.76 | 0.73 |
| **Secondary PL task accuracy** | Standalone | 81 | 0.51 (0.22) | 0 | 0.95 | 0.02 | -0.51 |
| With non-symbolic primary task | 81 | 0.51 (0.21) | 0 | 1 | -0.05 | -0.17 |
| With symbolic primary task | 81 | 0.50 (0.24) | 0.05 | 1 | 0.28 | -0.89 |
| With cross-modal primary task | 81 | 0.51 (0.22) | 0 | 0.9 | -0.23 | -0.69 |
| **Secondary VSSP task accuracy** | Standalone | 81 | 0.59 (0.60) | 0 | 1 | -0.61 | 0.99 |
| With non-symbolic primary task | 81 | 0.40 (0.20) | 0 | 0.95 | 0.18 | -0.36 |
| With symbolic primary task | 81 | 0.47 (0.21) | 0 | 1 | 0.21 | -0.4 |
|  | With cross-modal primary task | 81 | 0.38 (0.23) | 0 | 0.85 | 0.22 | -0.8 |

Note. PL = Phonological, VSSP = Visuospatial

**Table 2.**

*Descriptive statistics for primary and secondary task performance (accuracy and RT in seconds) in all experimental conditions for small and large quantities separately.*

|  |  |  |  |
| --- | --- | --- | --- |
|  |  | **Small Quantities** | **Large Quantities** |
|  |  | **N** | **Mean (SD)** | **Min** | **Max** | **Skewness** | **Kurtosis** | **N** | **Mean (SD)** | **Min** | **Max** | **Skewness** | **Kurtosis** |
| **Non-Symbolic primary task standalone** | Accuracy | 80 | 0.96 (0.04) | 0.79 | 1 | -2.49 | 4.45 | 80 | 0.88 (0.07) | 0.70 | 1 | -0.57 | -0.23 |
| Median RT | 79 | 0.51 (0.08) | 0.38 | 0.73 | 0.57 | -0.15 | 79 | 0.59 (0.13) | 0.35 | 1 | 0.82 | 0.87 |
| **Non-Symbolic primary task with PL secondary task** | Accuracy | 80 | 0.97 (0.03) | 0.85 | 1 | -1.76 | 3.34 | 80 | 0.89 (0.07) | 0.72 | 1 | -0.66 | -0.12 |
| Median RT | 79 | 0.53 (0.09) | 0.37 | 0.79 | 0.77 | 0.18 | 79 | 0.60 (0.13) | 0.34 | 0.97 | 0.94 | 0.80 |
| **Non-Symbolic primary task with VSSP secondary task** | Accuracy | 80 | 0.96 (0.04) | 0.81 | 1 | -1.77 | 4.45 | 80 | 0.88 (0.07) | 0.63 | 1 | -1.07 | 1.61 |
| Median RT | 79 | 0.53 (0.08) | 0.38 | 0.80 | 1.04 | 1.84 | 79 | 0.59 (0.11) | 0.37 | 0.88 | 0.83 | 0.56 |
| **Symbolic primary task standalone** | Accuracy | 79 | 0.98 (0.02) | 0.91 | 1 | -1.22 | 0.99 | 79 | 0.94 (0.05) | 0.75 | 1 | -1.03 | 1.47 |
| Median RT | 79 | 0.47 (0.07) | 0.37 | 0.71 | 0.79 | 0.35 | 79 | 0.50 (0.09) | 0.37 | 0.80 | 0.88 | 0.63 |
| **Symbolic primary task with PL secondary task** | Accuracy | 80 | 0.98 (0.02) | 0.94 | 1 | -0.91 | 1.23 | 81 | 0.95 (0.05) | 0.80 | 1 | -1.42 | 2.13 |
| Median RT | 79 | 0.49 (0.07) | 0.36 | 0.66 | 0.48 | -0.10 | 79 | 0.52 (0.09) | 0.38 | 0.77 | 0.54 | 0.09 |
| **Symbolic primary task with VSSP secondary task** | Accuracy | 81 | 0.99 (0.02) | 0.89 | 1 | -2.17 | 7.52 | 80 | 0.95 (0.05) | 0.80 | 1 | -1.15 | 1.04 |
| Median RT | 79 | 0.48 (0.07) | 0.37 | 0.67 | 0.57 | 0.11 | 79 | 0.53 (0.08) | 0.39 | 0.73 | 0.43 | -0.34 |
| **Cross-modal primary task standalone** | Accuracy | 80 | 0.92 (0.05) | 0.78 | 1 | -0.61 | 6.00 | 80 | 0.89 (0.08) | 0.63 | 1 | -0.83 | 0.75 |
| Median RT | 79 | 0.61 (0.12) | 0.39 | 1.12 | 1.96 | 6.75 | 79 | 0.68 (0.14) | 0.41 | 1.20 | 1.43 | 3.04 |
| **Cross-modal primary task with PL secondary task** | Accuracy | 78 | 0.92 (0.05) | 0.75 | 1 | -0.99 | 1.11 | 81 | 0.86 (0.09) | 0.63 | 1 | -0.89 | 0.39 |
| Median RT | 79 | 0.62 (0.11) | 0.42 | 1.06 | 1.37 | 2.96 | 79 | 0.66 (0.14) | 0.42 | 1.26 | 1.98 | 5.82 |
| **Cross-modal primary task with VSSP secondary task** | Accuracy | 79 | 0.92 (0.05) | 0.75 | 1 | -1.20 | 1.95 | 80 | 0.87 (0.10) | 0.60 | 1 | -1.09 | 1.09 |
| Median RT | 79 | 0.62 (0.10) | 0.4 | 0.94 | 0.82 | 1.09 | 79 | 0.67 (0.11) | 0.43 | 1.00 | 0.80 | 0.87 |

Note. PL = Phonological, VSSP = Visuospatial

Subsequently, we ran the analyses addressing our three research questions as outlined in our detailed analysis plan (Appendix C). Note that for some bivariate comparisons the data did not exceed kurtosis limits once all relevant exclusions had been performed and therefore a t-test was performed even if the individual variables exceeded this limit.

***RQ1. Are Arabic symbols and non-symbolic representations accessed automatically or does access require the involvement of WM components?***

*RQ1a. Symbolic comparison primary task*

We ran 6 paired sample t-tests focusing on the symbolic comparison primary task (see Figure 3). The t-test comparing participants’ accuracy in symbolic comparison with and without PL interference demonstrated a significant difference, t(78) = 1.98, *p* = .05, *d* = 0.22, however, in the opposite direction to our expectations. Specifically, participants performed better with PL interference than without. We also found a significant difference, t(78) = 4.55, *p* < .001, *d* = 0.51, when comparing the corresponding median RT data although this time in the expected direction, namely participants were slower in the symbolic primary task with PL interference than the one without. Contrary to our predictions, the paired sample t-tests comparing symbolic performance with and without VSSP interference demonstrated significant differences both with the accuracy, t(78) = 2.75, *p* < .05, *d* = 0.31, and median RT data, t(78) = 4.23, *p* < .001, *d* = 0.48, with better accuracy but slower performance in the VSSP dual-task compared to the standalone condition. Finally, comparing the two interference conditions (i.e., symbolic primary task with PL dual-task vs symbolic primary task with VSSP dual-task), we found no difference either for accuracy, t(80) = 0.58, *p* = .57, *d* = 0.06, or median RT, t (78) = 0.09, *p* = .93*, d* = .01.

Turning now to secondary task performance (Figure 4), with the symbolic primary task we found no difference between participants’ accuracy on the PL secondary task between the standalone and the interference (dual-task) condition, t(80) = 0.59, *p =* .56, *d* = 0.07*,* but there was a significant difference between participants’ accuracy on the VSSP secondary task conducted standalone versus the symbolic dual-task condition, t(80) = 5.31, *p* < .001, *d* = 0.59, with participants performing better in the standalone VSSP compared to the symbolic dual-task condition.

*RQ1b. Non-symbolic comparison primary task*

We ran 6 paired sample t-tests focusing on the non-symbolic comparison primary task (Figure 3). As expected, the comparison between the non-symbolic task conducted with and without PL interference showed no significant difference in accuracy, t(78) = 1.00, *p* = .32, *d* = 0.11, however, surprisingly, there was a significant difference in the median RT data, t(78) = 3.27, *p* = .002, *d* = 0.37, with participants being slower in the interference compared to the standalone condition. The paired sample t-tests comparing non-symbolic with and without VSSP interference demonstrated no difference with the accuracy data, t(79) = 0.90, *p* = 0.37*, d* = 0.10, but a significant difference with the median RT data, t(78) = 2.22, *p* = 0.03*, d* = 0.25, with participants being slower in the VSSP interference compared to the standalone condition. Finally, performance in the two non-symbolic interference conditions (PL and VSSP) significantly differed, t(79) = 2.44, *p* = .02, *d* = 0.27, with participants performing better under PL than VSSP interference. Median RTs did not differ between these two conditions, t(78) = 1.48, *p* = .14, *d* = 0.17.

Turning our focus now to the secondary tasks (Figure 4), as expected, we found no difference between the PL secondary performed standalone versus the non-symbolic interference condition, t(80) = 0.11, *p* = 0.91, *d* = 0.01, however, once again participants’ accuracy significantly dropped in the VSSP secondary task in the non-symbolic interference versus the standalone condition, t(80) = 8.80, *p* < .001, *d* = 0.98.

***RQ2. Can adults translate between Arabic and non-symbolic representations automatically or does this require access to verbal representations?***

Once again, we ran 6 paired sample t-tests this time focusing on the cross-modal comparison primary task (Figure 3). The comparison between the cross-modal with and without PL interference revealed no significant difference in accuracy, t(77) = 1.64, *p* = .10, *d* = 0.19, or median RTs, V = 1566, *p* = .90, *d* = 0.04 (non-parametric Wilcoxon signed-rank). There was also no difference between the cross-modal with and without VSSP interference in accuracy, t(77) = 1.94, *p* = .056, *d* = 0.22, or median RT, V = 1597, *p* = .78, *d* = 0.01 (non-parametric Wilcoxon signed-rank). Further, we found no difference between the two interference conditions (PL and VSSP) in accuracy, t(77) = 0.67, *p* = .50, *d* = 0.08, or median RT, V = 1497, *p* = .69, *d* = 0.04(non-parametric Wilcoxon signed-rank).

The expected involvement of WM in cross-modal comparison was found when examining participants’ secondary task performance (Figure 4). Specifically, we found that participants were more accurate in the VSSP secondary task when performed standalone compared to its cross-modal dual-task counterpart, t(80) = 10.34, *p* < .001, *d* = 1.15. However, there was no significant difference between the standalone PL secondary and its cross-modal dual-task counterpart t(80) = 0.03, *p* = .97, *d* = 0.003.

****

Figure 3.*Accuracy and median RT (seconds) on the three primary tasks (symbolic, non-symbolic and cross-modal) for standalone and interference (dual-task) conditions. Paired comparison results are indicated by \* p < .05, \*\* p < .001.*

****

**Figure 4.** Accuracy on the PL and VSSP secondary tasks conducted standalone versus dual-task conditions (symbolic, non-symbolic and cross-modal). Paired comparison results are indicated by \* *p < .05, \*\* p < .001.*

***RQ3: Does the processing of numerical information differ for small and large quantities?***

*RQ3a. Symbolic and cross-modal comparison*

For symbolic and cross-modal comparison, we had hypothesised PL involvement and no difference in the effect of PL interference for small and large numbers. In symbolic comparison, paired samples t-tests comparing small-number trials with and without PL interference demonstrated a significant difference in median RT, t(78) = 4.70, *p* < .001, *d* = 0.53, but not in accuracy, t(77) = 0.69, *p* = .49, *d* = 0.08. As expected, participants were slower in the small-number trials in the PL interference compared to standalone condition. For large quantities, the same comparison revealed a significant difference between with and without PL interference in both accuracy, t(78) = 2.35, *p* = .02, *d* = 0.26, and median RT, t(78) = 3.49, *p* < .001, *d* = 0.39. For large-number trials, participants performed better but slower in the interference condition compared to their standalone counterpart.

For cross-modal comparison, we found no difference for the small-quantity trials with and without PL interference (accuracy: t(76) = 0.42, *p* = .68, *d* = 0.05; median RT: V = 1332, *p =* .30, *d* = 0.13 - non-parametric Wilcoxon signed-rank). However, we found a significant difference in the corresponding large-quantity trials with both accuracy (t(79) = 2.72, *p* = .008, *d* = 0.30) and median RT (V = 1952, *p* = .04, *d* = 0.16 - non-parametric Wilcoxon signed-rank). Participants performed better in the cross-modal large-quantity trials without PL interference but faster in the corresponding trials with PL interference.

*RQ3b. Non-symbolic comparison*

For the case of non-symbolic comparison, we had hypothesised a difference in the effects of VSSP interference between small and large quantities. For small quantities, paired samples t-tests comparing trials with and without VSSP interference demonstrated no difference in accuracy (V = 1193.5, *p* = .20, *d* = 0.08 (non-parametric Wilcoxon signed-rank), but a significant difference in median RT, namely participants were slower in the dual-task small quantity trials compared to the standalone small quantity trials (t(78) = 2.70, *p* = .009, *d* = 0.30). However, we found no difference between the large quantities with and without VSSP interference (Accuracy: *t*(79) = 0.71, *p* = 0.47, *d* = 0.08, Median RT: *t*(78) = 0.17, *p* = .87, *d* = 0.02. Figure 5 summarises all the planned comparisons (black arrows) addressing our research questions as outlined above.

**

**Figure 5.** *Overview of all significant comparisons (ps < .05) across the experimental conditions of the 3 number comparison conditions (symbolic, non-symbolic and cross-modal). Black arrows depict the preregistered comparisons, and grey arrows the exploratory comparisons (see Appendix D).* Arrows facing upwards indicate better accuracy or faster speed in the interference condition (dual-task) whereas arrows facing downwards indicate reduced accuracy or slower speed in the interference condition (dual-task).

**Discussion**

Basic number processing is defined as the ability to understand, estimate or discriminate between numerical magnitudes (Holloway & Ansari, 2009; Sokolowski et al., 2022). The present study examined whether adults process numbers presented in different modalities and translate between them automatically or if doing so requires Working Memory (WM) resources. Participants conducted non-symbolic, symbolic and cross-modal comparison tasks with and without phonological (PL) and visuospatial (VSSP) WM interference. Interference was implemented by asking participants to perform the numerical comparison tasks while also undertaking a WM task where they had to retain either letters that they heard (PL) or visuospatial patterns that they saw (VSSP) and then recall them backwards. As these secondary tasks required the manipulation of the stored elements, they also involved the CE component of WM, i.e. the limited-capacity attentional system overseeing and coordinating the activities of the PL and VSSP (Baddeley, 2001; Hitch, Allen, & Baddeley, 2024; Repovš & Baddeley 2006). The PL and VSSP WM interference tasks were also completed as standalone tasks. Results demonstrated for the first time that all three types of comparison necessitate WM resources. Symbolic comparison necessitated VSSP WM, i.e. the VSSP component of WM in combination with the CE. Surprisingly, in this case, participants’ accuracy improved under both interference conditions, revealing the positive effects that embedding executive function challenges can have on simple and familiar numerical tasks. Non-symbolic comparison necessitated VSSP WM but also PL albeit to a lesser extent. Under VSSP interference, performance was reduced in both the primary and the secondary task, whereas under PL interference only in the primary and significantly less than the VSSP. Lastly, cross-modal comparison primarily taxed VSSP WM processes. Also, we observed different interference effects in processing and translating numerical representations for smaller (1 - 4) and larger (5 – 9) numerosities, which were examined separately. Overall, our findings provide clear evidence that basic number processing in adults is not automatic; it necessitates WM resources. Adults require WM resources for both simply processing numerical representations (non-symbolic and symbolic) as well as for translating between them. Our results also demonstrate the flexible and dynamic role that WM plays in basic number processing, aiding in the adaptive utilisation of phonological and visuospatial strategies when processing numerical representations. Below we discuss our findings in more detail and the respective theoretical and methodological implications.

*Symbolic magnitude comparison*

We start learning and understanding symbols from a very young age and continue to develop our symbolic number skills throughout our lives. It is unsurprising, therefore, that it is intuitively assumed that by adulthood we process symbolic numbers automatically with little conscious effort. The typical symbolic comparison task used to assess basic number processing, where participants see two digits on the screen and must decide which is larger as quickly and as accurately as possible, can be considered effortless for adults. Our dual-task paradigm revealed that WM is actively involved in the flexible processing of numerical representations when undertaking such a symbolic comparison task.

Surprisingly, participants’ accuracy improved instead of dropping under both types of the implemented active CE interference, i.e. with both the PL and VSSP WM interference. This was triggered by the interference conditions. At the same time, participants were slower under both interference conditions compared to the one without. Perhaps with WM interference, which made the task more challenging and potentially less tedious, participants’ focus on the primary task improved, leading to improved accuracy and slower responses. Conversely, the challenge induced by the interference conditions may have slowed participants down, leading to improved performance. Our exploratory analyses (Appendix D, Table D3) showed a significant but small speed-accuracy trade-off (r = 0.20) for the condition with PL interference, but not for the condition with VSSP interference nor the condition without interference. However, these speed-accuracy correlation coefficients did not differ significantly, which may suggest no change in participants’ speed-accuracy trade-off across the conditions. Thus, a speed-accuracy trade-off on its own does not appear to explain the finding of improved accuracy in interference conditions. We should highlight here though that this is only a provisional assumption based on exploratory analyses for which we had not conducted a power analysis, and should be therefore considered tentative. Nevertheless, it is clear that accuracy improved due to the interference conditions, therefore there must be WM involvement. The fact that we found no difference between the two interference conditions (in both accuracy and RT) indicates that both had a similar effect on symbolic comparison. As expected, results were similar when splitting the trials based on the size of the number, i.e. there was no difference in the effect of interference for trials on symbolic comparison with small or large quantities.

The role of WM in symbolic comparison became clearer when examining the trade-off between the primary and the secondary tasks. Participants’ performance dropped in the VSSP secondary task, but not the PL secondary task. Similarly, Van Dijk, Gevers and Fias (2009) and Herrera (2008), who implemented verbal and visuospatial load on adults’ symbolic magnitude comparison processing, found performance was affected only under the VSSP, not the PL load. In those studies, the WM load only required the storage of elements of information during the dual conditions, whereas the secondary tasks used in the present study also necessitated processing the given visuospatial elements. Thus, our study extends past findings by demonstrating that symbolic comparison necessitates not just VSSP storage but also VSSP processing, i.e. the combination of the VSSP component of WM with the CE. This suggests that participants use VSSP strategies to solve this type of numerical comparison task, for example by using nonsymbolic or mental number line representations for the symbolic quantities. These results are in line with the triple-code model (Dehaene, 1992), i.e. the involvement of VSSP WM may reflect the involvement of the analogue code in symbol-symbol comparison, whereas the lack of evidence for PL WM involvement suggests that the verbal code is not necessary for this type of numerical comparison. Notably, the triple-code model (Dehaene, 1992) does not mention the potential role that WM may play in mentally representing and processing numerical representations. In Dehaene’s (1992) review, WM is only mentioned concerning its role in arithmetic. The present study extends this model by evidencing the key role that WM plays in processing and translating between the different types of numerical representations.

Further, our findings for symbolic comparison suggest that the fact that a task requires little cognitive effort does not mean that it is processed automatically. On the contrary, we observed widespread interference effects in this primary task (see Figure 5). Embedding executive function challenges in easy tasks can potentially enhance performance. Indeed, in cognitive psychology, there have been other instances where under certain conditions participants’ performance improved under WM load conditions (Giammarco et al., 2015). For example, Makovski et al. (2008) demonstrated that orienting attention in visual WM under dual-task conditions can lead to improved performance by reducing interference from memory probes. Thus, it appears that attentional mechanisms can mitigate WM interference under certain conditions. Taken together, these findings reveal the complexity of cognitive processes and the potential for adaptive mechanisms to enhance performance even in the presence of interference (Giammarco et al., 2015).

*Non-symbolic magnitude comparison*

In the non-symbolic magnitude comparison task, performance was affected under both WM load interference conditions, with participants being slower under both conditions. Importantly, this time there was a significant difference in the impact of the two interference conditions; performance with the PL interference task was higher than its VSSP counterpart. Finally, participants’ performance again dropped in the VSSP secondary task, but not in the PL secondary task. Taken together, these findings provide clear evidence that both subcomponents of WM, the VSSP and the PL are needed in combination with the CE when processing non-symbolic quantities. In 2014, Xenidou-Dervou, Van Lieshout and Van der Schoot conducted a dual-task study with young children where they implemented phonological, visual, spatial and CE interference on a non-symbolic approximate arithmetic task, that is a task where participants see two dot arrays dropping inside a box on one side of the screen before having to make the judgment on whether the sum of the arrays was larger or smaller than a third dot array. They found that performance dropped only under the CE interference condition, where participants had to process phonological information while undertaking the primary task. We find this PL-CE combination to be crucial for simple non-symbolic magnitude comparison in adults too. Notably, in Xenidou-Dervou et al. (2014), the non-symbolic task entailed only large quantities (6 up to 70 dots) and the visual and spatial secondary tasks only necessitated storage of the corresponding elements. Thus, the involvement of the VSSP-CE combination, i.e. VSSP WM, was not directly examined, although it can be argued that to an extent dual tasks inherently necessitate executive functions (such as switching between two tasks) (Chen & Bailey, 2021). The findings of the present study clearly demonstrate that visuospatial processing is not only necessary but also more important than phonological processing for non-symbolic magnitude comparison.

Another interesting finding arose when considering the size of the quantities involved in the trials. Contrary to our hypotheses, when considering only primary task performance, the VSSP interference was observed only in the small-quantity trials, not the large (although the involvement of the VSSP overall is still evident from the secondary-task results). Given the unexpected involvement of the PL in non-symbolic comparison, we opted to conduct similar exploratory analyses splitting the trials on the basis of the size of the quantities under the PL interference condition too and found the PL to be involved again only in the small-quantity trials. This pattern of results in the small-quantity trials is the same as that observed for the symbolic comparison task, tentatively supporting the assumption that small non-symbolic quantities may be processed similarly to symbolic representations (Hutchison et al., 2019).

*Cross-modal magnitude comparison*

 For the cross-modal primary task, we had hypothesised that the translation between symbolic and non-symbolic representations would require access to verbal representations and therefore the PL. Surprisingly, however, participants’ performance on the cross-modal comparison primary task was not affected by either WM interference conditions. Instead, their performance dropped once again in the VSSP secondary task, but not the PL secondary task, revealing yet again the crucial role that visuospatial processing plays not only in processing numerical representations but also in translating between them. The finding that VSSP, but not PL, WM is necessary for cross-modal comparison addresses an open question in the triple-code model (Dehaene, 1992); it suggests that translating from non-symbolic to symbolic representations is possible without accessing the verbal code.

Interestingly, we found some evidence of potential involvement of the PL only for the large-quantity trials, which suggests that translation between symbolic and non-symbolic representations may require the processing of verbal representations only for large quantities. However, these results are not conclusive: Although accuracy dropped in the cross-modal large-quantity trials, unexpectedly corresponding RT dropped too (i.e. participants were faster under PL WM load) and the difference between the two interference conditions was not significant, indicating overall additional WM load from completing two tasks in these trials, i.e. CE involvement.

*Overview*

The present study’s findings extend our understanding of the causal mechanisms that drive basic number processing skills and the fundamental role that WM processing plays herein. We found that all three types of comparison employ WM resources: Symbolic and cross-modal involve VSSP WM, whereas non-symbolic comparison involves both VSSP and PL WM. We had hypothesised that numerical processing would necessitate not just the mere storage of numerical representations in WM but also their processing, therefore we used secondary tasks that would tax the CE component as well as the storage components of WM. Indeed, we found that although the different forms of numerical processing may differ on the type of subcomponent that they may employ (PL or VSSP), therefore reflecting the type of strategies an individual may use to undertake the task at hand, i.e. phonological or visuospatial, we see that all forms of numerical processing necessitated the CE one way or another, i.e. the processing of the different types of numerical representations.

Interestingly, and unexpectedly, we found that our interference conditions affected the harder task of translating between modalities, i.e. cross-modal comparison, in fewer performance outcomes compared to the easier symbolic and non-symbolic comparison tasks. It is well known that translating between numerical representations is harder than non-symbolic processing and this in turn is harder than symbolic processing (Lyons et al., 2012). We confirmed this to be the case in our data too (see exploratory analyses in Appendix D). Inspecting Figure 5, which provides an overview of all the performance outcomes that were affected (or not) by the WM load conditions, one notices the most effects in symbolic comparison and the least in cross-modal comparison. This is the opposite of what one may have intuitively expected as we tend to think in terms of the extent of reduced performance due to WM load in a particular component, not across the breadth of available cognitive resources. What this bird’s eye view of the overall results demonstrates is that when engaging in more difficult tasks such as translating between numerical representations, individuals may use their WM resources less flexibly. This phenomenon is typically supported by research showing that extensive mental effort can lead to the depletion of our limited WM resources, resulting in decreased performance compared to tasks requiring less mental effort (e.g., Chen, Castro-Alonso, Paas, & Sweller, 2018). Our findings extend this even to cross-modal number comparison, which although may be more taxing compared to the other notation-specific comparison tasks used (symbolic or non-symbolic), it is not that arduous in relative terms. It is in fact considered a very basic number-processing skill for adults. Accumulating these findings and assumptions, perhaps we should not view WM only in terms of a bottleneck – a limited-capacity mental space perspective that considers harder tasks taking up more cognitive resources, but rather that harder tasks lead to less flexible/adaptable usage of an individual’s available WM resources.

*Limitations and Future Directions*

The dual-task paradigm is a robust experimental design, which allows for causal conclusions to be drawn regarding underlying mechanisms (Raghubar et al., 2009). The present study has many methodological strengths; we used active WM interference conditions which allowed us to examine the trade-off between primary and secondary tasks, differential response modalities for primary and secondary tasks, powered our study adequately, pre-registered methods and analytical approach. However, our approach is not free of limitations. Firstly, our design is limited when it comes to understanding the functions of the CE, which we found to be a crucial component (Chen & Bailey, 2021). Future studies should strive to unravel the specific functions of the CE involved in number processing and the conditions under which attentional mechanisms and executive function challenges can mitigate WM interferences and even enhance performance. Secondly, although the secondary tasks that we used can be seen as a strength as we extend previous studies, which had only used simple WM load conditions that did not require processing of information in one’s WM, future studies should directly contrast the effects of storage vs processing. Given the prominent role of the VSSP component in all forms of basic number processing in our study, future research should zoom into this component and examine the role of the visual and spatial subcomponents separately (Baddeley, 2003; Xenidou-Dervou et al., 2014). Finally, our study allows limited conclusions regarding the comparison of the effects of WM load on small or large numbers/quantities. That is because WM load was imposed before and after the presentation of eight trials of the primary task, where both small and large numbers could have been shown. Future research should use a blocked design where the WM interference is imposed separately for small and large trials for a more robust comparison.

Finally, beyond the theoretical and methodological implications, this study’s findings also generate an interesting practical implication: Even a simple task such as comparing small numbers or quantities, may be more taxing to one’s limited WM resources than intuitively expected. From an educational perspective, future research should examine how taxing numerical processing is for children, who are still in the process of learning and developing their numerical skills, and the conditions under which increasing or decreasing the WM demands of a numerical task may benefit or hinder performance and learning.

*Concluding remarks*

How we process numerical representations is a fundamental question in the field of numerical cognition. The present study’s findings extend our theoretical understanding by revealing that although numerical representations may be activatedautomatically (e.g. Besner & Coltheart, 1979; Reike & Schwarz, 2017; Xenidou-Dervou et al., 2015), when processing is needed WM resources are employed. WM resources are required both for simply processing numerical representations (non-symbolic and symbolic comparison) as well as for translating between them (cross-modal comparison). The flexible use of visuospatial strategies seems to be most crucial for adults’ basic number processing, confirming the strong link between numbers and space in numerical cognition (Cipora et al., 2018).

**Conflict of interest disclosure.**

The authors of this article declare that they have no financial conflict of interest with the content of this article.

**References**

Alloway, T. P., & Alloway, R. G. (2013). Working memory across the lifespan: A cross-sectional approach. *Journal of Cognitive Psychology*, *25*(1), 84-93. https://doi.org/10.1080/20445911.2012.748027

Baddeley, A. (2000). The episodic buffer: A new component of working memory? *Trends in Cognitive Sciences*, *4*(11), 417–423. [https://doi.org/10.1016/S1364-6613(00)01538-2](https://doi.org/10.1016/S1364-6613%2800%2901538-2)

Baddeley, A. D. (2001). Is working memory still working? *American psychologist, 56*(11), 851.

Baddeley, A. (2003). Working memory and language: an overview. Journal of Communication Disorder, 36, 189–208.

Baddeley, A. D. (2010). Working memory. *Current Biology*, *20*(4), R136–R140.

Baddeley, A. D., & Hitch, G. (1974). Working memory. In *Psychology of learning and motivation* (Vol. 8, pp. 47–89). Elsevier.

Barner, D. (2018). *Symbols explain perception*. 1–59.

Batchelor, S., Keeble, S., & Gilmore, C. (2015). Magnitude representations and counting skills in preschool children. *Mathematical Thinking and Learning*, *17*(2–3), 116–135.

Bernoussi, M., & Khomsi, A. (1997). Transcoding numbers: A developmental approach of comprehension. *European Journal of Psychology of Education*, *12*(3), 293–304. https://doi.org/10.1007/BF03172877

Besner, D., & Coltheart, M. (1979). Ideographic and alphabetic processing in skilled reading of English. *Neuropsychologia*, *17*(5), 467–472. https://doi.org/10.1016/0028-3932(79)90053-8

Blau, V., Reithler, J., van Atteveldt, N., Seitz, J., Gerretsen, P., Goebel, R., & Blomert, L. (2010). Deviant processing of letters and speech sounds as proximate cause of reading failure: A functional magnetic resonance imaging study of dyslexic children. *Brain*, *133*(3), 868–879. https://doi.org/10.1093/brain/awp308

Brankaer, C., Ghesquière, P., & De Smedt, B. (2014). Children’s mapping between non-symbolic and symbolic numerical magnitudes and its association with timed and untimed tests of mathematics achievement. *PLoS ONE*, *9*(4), e93565. https://doi.org/10.1371/journal.pone.0093565

Chen, E. H., & Bailey, D. H. (2021). Dual-task studies of working memory and arithmetic performance: A meta-analysis. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 47(2),* 220.

Chen, O., Castro-Alonso, J. C., Paas, F., & Sweller, J. (2018). Extending cognitive load theory to incorporate working memory resource depletion: Evidence from the spacing effect. *Educational Psychology Review, 30*, 483-501. DOI 10.1007/s10648-017-9426-2.

Cipora, K., Schroeder, P. A., Soltanlou, M., & Nuerk, H. C. (2018). More space, better mathematics: Is space a powerful tool or a cornerstone for understanding arithmetic?. Visualizing mathematics: The role of spatial reasoning in mathematical thought, 77-116.

Dehaene, S. (1992). Varieties of numerical abilities. *Cognition*, *44*(1), 1–42.

Dehaene, S., Bossini, S., & Giraux, P. (1993). The mental representation of parity and number magnitude. Journal of Experimental Psychology: General, 122, 371–396.

Department for Business, Innovation and Skills. (2011). *2011 skills for life survey: Headline findings* (p. 25).

Diamond, A. (2013). Executive functions. *Annual Review of Psychology*, *64*(1), 135–168. https://doi.org/10.1146/annurev-psych-113011-143750

Dietrich, J. F., Huber, S., & Nuerk, H.-C. (2015). Methodological aspects to be considered when measuring the Approximate Number System (ANS): A research review. *Frontiers in Psychology*, *6*. https://doi.org/10.3389/fpsyg.2015.00295

Ding, Y., Liu, R.-D., Xu, L., Wang, J., & Zhang, D. (2017). Working memory load and automaticity in relation to mental multiplication. *The Journal of Educational Research*, *110*(5), 554–564. https://doi.org/10.1080/00220671.2016.1149794

Duncan, G. J., Dowsett, C. J., Claessens, A., Magnuson, K., Huston, A. C., Klebanov, P., Pagani, L. S., Feinstein, L., Engel, M., Brooks-Gunn, J., Sexton, H., Duckworth, K., & Japel, C. (2007). School readiness and later achievement. *Developmental Psychology*, *43*(6), 1428–1446. https://doi.org/10.1037/0012-1649.43.6.1428

Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G\*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, *41*(4), 1149–1160. <https://doi.org/10.3758/BRM.41.4.1149>

Field, A. (2016). Discovering statistics using IBM SPSS statistics. SAGE publications.

Feigenson, L., Dehaene, S., & Spelke, E. (2004). Core systems of number. *Trends in Cognitive Sciences*, *8*(7), 307–314. https://doi.org/10.1016/j.tics.2004.05.002

Friso-van den Bos, I., van der Ven, S. H. G., Kroesbergen, E. H., & van Luit, J. E. H. (2013). Working memory and mathematics in primary school children: A meta-analysis. *Educational Research Review*, *10*, 29–44. https://doi.org/10.1016/j.edurev.2013.05.003

Gallistel, C. R., & Gelman, R. (2000). Non-verbal numerical cognition: From reals to integers. *Trends in cognitive sciences*, *4*(2), 59-65. [https://doi.org/10.1016/S1364-6613(99)01424-2](https://doi.org/10.1016/S1364-6613%2899%2901424-2)

Geary, D. C., Hamson, C. O., & Hoard, M. K. (2000). Numerical and arithmetical cognition: A longitudinal study of process and concept deficits in children with learning disability. *Journal of Experimental Child Psychology*, *77*(3), 236–263. https://doi.org/10.1006/jecp.2000.2561

Giammarco, M., Thomson, S. J., & Watter, S. (2016). Dual-task backward compatibility effects are episodically mediated. Attention, Perception, & Psychophysics, 78, 520-54. https://doi.org/10.3758/s13414-015-0998-y

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. <https://doi.org/10.1371/journal.pone.0067374>

Gilmore, C., Keeble, S., Richardson, S., & Cragg, L. (2015). The role of cognitive inhibition in different components of arithmetic. *Zdm*, *47*(5), 771-782. DOI 10.1007/s11858-014-0659-y

Göbel, S. M., Watson, S., Lervag, A., & Hulme, C. (2014). Children’s arithmetic development: It is number knowledge, not the approximate number sense, that counts. *Psychological Science*, *25*(3), 789–798. <https://doi.org/10.1177/0956797613516471>

Guan, D., Ai, J., Gao, Y., Li, H., Huang, B., & Si, J. (2021). Non-symbolic representation is modulated by math anxiety and cognitive inhibition while symbolic representation not. *Psychological Research*, *85*(4), 1662-1672. https://doi.org/10.1007/s00426-020-01356-7

Gunderson, E. A., Spaepen, E., & Levine, S. C. (2015). Approximate number word knowledge before the cardinal principle. *Journal of Experimental Child Psychology*, *130*, 35–55. https://doi.org/10.1016/j.jecp.2014.09.008

Habermann, S., Donlan, C., Göbel, S. M., & Hulme, C. (2020). The critical role of Arabic numeral knowledge as a longitudinal predictor of arithmetic development. *Journal of Experimental Child Psychology*, *193*, 104794. https://doi.org/10.1016/j.jecp.2019.104794

Halberda, J., & Feigenson, L. (2008). Developmental change in the acuity of the ‘number sense’: The approximate number system in 3-, 4-, 5-, and 6-year-olds and adults. *Developmental Psychology*, *44*(5), 1457–1465. https://doi.org/10.1037/a0012682

Halberda, J., Ly, R., Wilmer, J. B., Naiman, D. Q., & Germine, L. (2012). Number sense across the lifespan as revealed by a massive Internet-based sample. *Proceedings of the National Academy of Sciences*, *109*(28), 11116–11120. <https://doi.org/10.1073/pnas.1200196109>

Herrera, A., Macizo, P., & Semenza, C. (2008). The role of working memory in the association between number magnitude and space. *Acta Psychologica*, *128*(2), 225-237. doi:10.1016/j.actpsy.2008.01.002

Hitch, G. J., Allen, R. J., & Baddeley, A. D. (2024). The multicomponent model of working memory fifty years on. *Quarterly Journal of Experimental Psychology*, 17470218241290909.

Holloway, I. D., & Ansari, D. (2009). Mapping numerical magnitudes onto symbols: The numerical distance effect and individual differences in children’s mathematics achievement. *Journal of Experimental Child Psychology*, *103*(1), 17–29. https://doi.org/10.1016/j.jecp.2008.04.001

Holloway, I. D., van Atteveldt, N., Blomert, L., & Ansari, D. (2015). Orthographic dependency in the neural correlates of reading: Evidence from audiovisual integration in English readers. *Cerebral Cortex*, *25*(6), 1544–1553. https://doi.org/10.1093/cercor/bht347

Hutchison, J. E., Ansari, D., Zheng, S., De Jesus, S., & Lyons, I. M. (2019). The relation between subitizable symbolic and non‐symbolic number processing over the kindergarten school year. *Developmental Science*. https://doi.org/10.1111/desc.12884

Izard, V., & Dehaene, S. (2008). Calibrating the mental number line. *Cognition*, *106*, 1221–1247.

Jiménez Lira, C., Carver, M., Douglas, H., & LeFevre, J.-A. (2017). The integration of symbolic and non-symbolic representations of exact quantity in preschool children. *Cognition*, *166*, 382–397. https://doi.org/10.1016/j.cognition.2017.05.033

Kessels, R. P., Van Zandvoort, M. J., Postma, A., Kappelle, L. J., & De Haan, E. H. (2000). The Corsi block-tapping task: Standardization and normative data. *Applied Neuropsychology*, *7*(4), 252–258.

Kline, R.B. (2011). *Principles and Practice of Structural Equation Modeling.* (3rd ed.), Guilford, NY.

Koponen, T., Aunola, K., & Nurmi, J.-E. (2019). Verbal counting skill predicts later math performance and difficulties in middle school. *Contemporary Educational Psychology*, 101803. https://doi.org/10.1016/j.cedpsych.2019.101803

Lakens, D. (2022). Sample size justification. *Collabra: Psychology*, *8*(1), 33267. https://doi.org/10.1525/collabra.33267.

Libertus, M. E., Feigenson, L., & Halberda, J. (2013). Is approximate number precision a stable predictor of math ability? *Learning and Individual Differences*, *25*, 126–133. https://doi.org/10.1016/j.lindif.2013.02.001

Lyons, I. M., Ansari, D., & Beilock, S. L. (2012). Symbolic estrangement: Evidence against a strong association between numerical symbols and the quantities they represent. *Journal of Experimental Psychology: General*, *141*(4), 635–641. https://doi.org/10.1037/a0027248

Maloney, E. A., Barr, N., Risko, E. F., & Fugelsang, J. A. (2019). Verbal working memory load dissociates common indices of the numerical distance effect: Implications for the study of numerical cognition. *Journal of Numerical Cognition*, *5*(3), 337–357. <https://doi.org/10.5964/jnc.v5i3.155>

Makovski, T., Sussman, R., & Jiang, Y. V. (2008). Orienting attention in visual working memory reduces interference from memory probes. Journal of Experimental Psychology: Learning, Memory, and Cognition, 34(2), 369. DOI: 10.1037/0278-7393.34.2.369

Moeller, K., Pixner, S., Zuber, J., Kaufmann, L., & Nuerk, H.-C. (2011). Early place-value understanding as a precursor for later arithmetic performance—A longitudinal study on numerical development. *Research in Developmental Disabilities*, *32*(5), 1837–1851. https://doi.org/10.1016/j.ridd.2011.03.012

Monaco, M., Costa, A., Caltagirone, C., & Carlesimo, G. A. (2013). Forward and backward span for verbal and visuo-spatial data: Standardization and normative data from an Italian adult population. *Neurological Sciences*, *34*(5), 749–754. https://doi.org/10.1007/s10072-012-1130-x

Mundy, E., & Gilmore, C. (2009). Children’s mapping between symbolic and nonsymbolic representations of number. *Journal of Experimental Child Psychology*, *103*(4), 490–502. <https://doi.org/10.1016/j.jecp.2009.02.003>

Norris, J. E., Clayton, S., Gilmore, C., Inglis, M., & Castronovo, J. (2019). The measurement of approximate number system acuity across the lifespan is compromised by congruency effects. *Quarterly Journal of Experimental Psychology*, *72*(5), 1037-1046. <https://doi.org/10.1177/1747021818779020>

O’Brien, T. (2014). *Exploring the Mappings between the Symbolic and Approximate Number Systems*. Loughborough University.

Odic, D., Le Corre, M., & Halberda, J. (2015). Children’s mappings between number words and the approximate number system. *Cognition*, *138*, 102–121. https://doi.org/10.1016/j.cognition.2015.01.008

Orrantia, J., Muñez, D., Matilla, L., Sanchez, R., San Romualdo, S., & Verschaffel, L. (2019). Disentangling the mechanisms of symbolic number processing in adults’ mathematics and arithmetic achievement. *Cognitive Science*, *43*(1). https://doi.org/10.1111/cogs.12711

Östergren, R., & Träff, U. (2013). Early number knowledge and cognitive ability affect early arithmetic ability. *Journal of Experimental Child Psychology*, *115*(3), 405–421. https://doi.org/10.1016/j.jecp.2013.03.007

Purpura, D. J., Baroody, A. J., & Lonigan, C. J. (2013). The transition from informal to formal mathematical knowledge: Mediation by numeral knowledge. *Journal of Educational Psychology*, *105*(2), 453–464. https://doi.org/10.1037/a0031753

Purpura, D. J., & Ganley, C. M. (2014). Working memory and language: Skill-specific or domain-general relations to mathematics? *Journal of Experimental Child Psychology*, *122*, 104–121. <https://doi.org/10.1016/j.jecp.2013.12.009>

R Core Team (2021). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.

Raghubar, K. P., Barnes, M. A., & Hecht, S. A. (2010). Working memory and mathematics: A review of developmental, individual difference, and cognitive approaches. *Learning and Individual Differences*, *20*(2), 110–122. https://doi.org/10.1016/j.lindif.2009.10.005

Reike, D., & Schwarz, W. (2017). Exploring the origin of the number-size congruency effect: Sensitivity or response bias? *Attention, Perception, & Psychophysics*, *79*(2), 383–388. https://doi.org/10.3758/s13414-016-1267-4

Repovš, G., & Baddeley, A. (2006). The multi-component model of working memory: Explorations in experimental cognitive psychology. Neuroscience, 139(1), 5-21.

Revkin, S. K., Piazza, M., Izard, V., Cohen, L., & Dehaene, S. (2008). Does subitizing reflect numerical estimation? *Psychological Science*, *19*(6), 607–614. https://doi.org/10.1111/j.1467-9280.2008.02130.x

Reynvoet, B., & Brysbaert, M. (2004). Cross-notation number priming investigated at different stimulus onset asynchronies in parity and naming tasks. *Experimental Psychology*, *51*(2), 81–90. https://doi.org/10.1027/1618-3169.51.2.81

Ritchie, S. J., & Bates, T. C. (2013). Enduring links from childhood mathematics and reading achievement to adult socioeconomic status. *Psychological Science*, *24*(7), 1301–1308. https://doi.org/10.1177/0956797612466268

Schleifer, P., & Landerl, K. (2011). Subitizing and counting in typical and atypical development: Subitizing and counting. *Developmental Science*, *14*(2), 280–291. https://doi.org/10.1111/j.1467-7687.2010.00976.x

Schneider, M., Beeres, K., Coban, L., Merz, S., Susan Schmidt, S., Stricker, J., & De Smedt, B. (2017). Associations of non-symbolic and symbolic numerical magnitude processing with mathematical competence: A meta-analysis. *Developmental Science*, *20*(3), e12372. <https://doi.org/10.1111/desc.12372>

Sokolowski, H. M., Hawes, Z., Leibovich-Raveh, T., & Ansari, D. (2022). Number symbols are processed more automatically than non-symbolic numerical magnitudes: Findings from a Symbolic-Non-symbolic Stroop task. *Acta Psychologica, 228*, 103644.

Stock, P., Desoete, A., & Roeyers, H. (2009). Mastery of the counting principles in toddlers: A crucial step in the development of budding arithmetic abilities? *Learning and Individual Differences*, *19*(4), 419–422. https://doi.org/10.1016/j.lindif.2009.03.002

Sullivan, J., & Barner, D. (2013). How are number words mapped to approximate magnitudes? *Quarterly Journal of Experimental Psychology*, *66*(2), 389–402. <https://doi.org/10.1080/17470218.2012.715655>

van Dijck, J. P., Gevers, W., & Fias, W. (2009). Numbers are associated with different types of spatial information depending on the task. *Cognition*, *113*(2), 248-253. doi:10.1016/j.cognition.2009.08.005

van Marle, K., Chu, F. W., Li, Y., & Geary, D. C. (2014). Acuity of the approximate number system and preschoolers’ quantitative development. *Developmental Science*, *17*(4), 492–505. https://doi.org/10.1111/desc.12143

Vanbinst, K., Ceulemans, E., Peters, L., Ghesquière, P., & De Smedt, B. (2018). Developmental trajectories of children’s symbolic numerical magnitude processing skills and associated cognitive competencies. *Journal of Experimental Child Psychology*, *166*, 232–250. https://doi.org/10.1016/j.jecp.2017.08.008

vanMarle, K., Chu, F. W., Mou, Y., Seok, J. H., Rouder, J., & Geary, D. C. (2018). Attaching meaning to the number words: Contributions of the object tracking and approximate number systems. *Developmental Science*, *21*(1), e12495. https://doi.org/10.1111/desc.12495

Wang, J., Halberda, J., & Feigenson, L. (2020). Emergence of the link between the approximate number system and symbolic math ability. *Child Development*, *0*(0), 1–15.

Wender, K. F., & Rothkegel, R. (2000). Subitizing and its subprocesses. *Psychological Research*, *64*(2), 81–92. https://doi.org/10.1007/s004260000021

Wynn, K. (1992). Children’s acquisition of the number words and the counting system. *Cognitive Psychology*, *24*, 220–251.

Xenidou-Dervou, I., De Smedt, B., van der Schoot, M., & van Lieshout, E. C. D. M. (2013). Individual differences in kindergarten math achievement: The integrative roles of approximation skills and working memory. *Learning and Individual Differences*, *28*, 119–129. https://doi.org/10.1016/j.lindif.2013.09.012

Xenidou-Dervou, I., Gilmore, C., van der Schoot, M., & van Lieshout, E. C. D. M. (2015). The developmental onset of symbolic approximation: Beyond nonsymbolic representations, the language of numbers matters. *Frontiers in Psychology*, *6*. https://doi.org/10.3389/fpsyg.2015.00487

Xenidou-Dervou, I., van der Schoot, M., & van Lieshout, E. C. D. M. (2015). Working memory and number line representations in single-digit addition: Approximate versus exact, nonsymbolic versus symbolic. *Quarterly Journal of Experimental Psychology*, *68*(6), 1148–1167. <https://doi.org/10.1080/17470218.2014.977303>

Xenidou-Dervou, I., van Lieshout, E.C.D.M., & van der Schoot, M. (2014). Working memory in non-symbolic approximate arithmetic processing: a dual-task study with preschoolers. *Cognitive Science, 38(1),* pp.101-127, ISSN: 0364-0213. DOI: 10.1111/cogs.12053.

Yang, X., Zhang, X., Huo, S., & Zhang, Y. (2020). Differential contributions of cognitive precursors to symbolic versus non-symbolic numeracy in young Chinese children. *Early Childhood Research Quarterly*, *53*, 208–216. https://doi.org/10.1016/j.ecresq.2020.04.003

Zuber, J., Pixner, S., Moeller, K., & Nuerk, H.-C. (2009). On the language specificity of basic number processing: Transcoding in a language with inversion and its relation to working memory capacity. *Journal of Experimental Child Psychology*, *102*(1), 60–77. https://doi.org/10.1016/j.jecp.2008.04.003

**Acknowledgements**

We would like to thank all the participants, who took part in this study. Further, we thank Krzysztof Cipora and Matthew Inglis for the insightful discussions on our findings and the PCI Registered Reports editor and reviewers for their constructive feedback and rigorous reviews. This work was partially supported by the UKRI Economic and Social Research Council [grant number ES/W002914/1] and a Royal Society Dorothy Hodgkin Fellowship (CG).

## Appendix A

***Table A.*** *Calculations of effect sizes.*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Primary condition | Analysis | Standalone mean (SD) | Minimal effect of interest (number of trials difference) | Minimal effect size of interest (Cohen’s d) | Resulting number of participants |
| Symbolic comparison | Primary accuracy | 99.6% (1.88) | 5 trials | 1.66 | 5 |
| Non-symbolic comparison | Primary accuracy | 99.7% (0.3) | 5 trials | 0.35 | 72  |
| Cross-modal comparison | Primary accuracy | 87.6% (9.38) | 5 trials | 0.33 | 81 |

|  |  |  |  |
| --- | --- | --- | --- |
| Primary condition | Analysis | Standalone mean (SD) | Minimal effect detectable with 81 participants (number of trials difference or increase in RT) |
| Symbolic comparison | Primary RT | 401ms (127) | 50ms |
|  | Secondary (PL) accuracy | 16 (4) | 1.3 sequences |
| Non-symbolic comparison | Primary RT | 499ms (141) | 50ms |
|  | Secondary (VSSP) accuracy | 12 (4) | 1.3 sequences |
| Cross-modal comparison | Primary RT | 799ms (244) | 80ms |
|  | Secondary (PL) accuracy | 16 (4) | 1.3 sequences |

**Appendix B**

### Power analysis for largest sample.



## Appendix C

***Detailed Analysis Plan***

Preliminary analyses

For each participant, mean accuracy will be calculated for the primary task and secondary task. Median RT (for correct trials only) will be calculated for the primary task conditions. Prior to conducting our main analysis (see table below), we will perform normality checks. Data will be plotted and skewness and kurtosis values will be examined. We expect some level of skew for the accuracy data (particularly for the Arabic digit condition) due to high accuracies in the standalone condition. Following recommendations (e.g., Kline et al., 2011) we will conduct non-parametric paired comparisons (Wilcoxon signed-rank) instead of parametric paired t-tests if skew is > |3|or kurtosis is > |4|. Outliers will be examined for performance on each task (i.e., primary and secondary tasks). Extreme outliers (> 3.29 SD, Field, 2016) will be removed from the analysis. All analyses were conducted in R.

To answer our three research questions, we will conduct all of the analyses described in the table below (analysis plan column). The alternative interpretations of the different potential outcomes are provided below. For all analyses described below, a “decrease in performance” refers to either a decrease in accuracy or an increase in reaction times between the stated conditions*.*

**Table C.** *Detailed analysis plan*

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Question** | **Hypothesis** | **Sampling plan** | **Analysis plan** | **Rationale for deciding the sensitivity of the test for confirming or disconfirming the hypothesis** | **Interpretation given different outcomes** | **Theory that could be shown wrong by the outcomes** |
| **RQ1a** Is the processing of Arabic digits automatic, or does it require the involvement of WM components?  | **RQ1a**. Processing of Arabic digits will require the involvement of the PL but not the VSSP. This means we will expect to see a difference in performance between the **symbolic primary task** in the PL dual-task condition compared to the VSSP dual-task condition and standalone condition, OR a difference between the **PL secondary task** in the dual-task condition when  | Based on the effect size calculations described above, the smallest effect size of interest for this RQ is 0.5.Using a power level of 90%, an alpha level of 0.05 and a minimum effect size of Cohen’s d = 0.5, the minimum sample required for this RQ is 36. | Primary tasks**Symbolic comparison (accuracy)**Paired sample t-test comparing: 1. Symbolic primary task with PL dual-task vs symbolic primary task standalone
2. Symbolic primary task with PL dual-task vs symbolic primary task with VSSP dual-task
3. Symbolic primary task with VSSP dual-task vs symbolic primary task standalone
 | The calculated effect sizes represent: a reduction in primary accuracy of 5 trials, or a reduction in secondary accuracy of 2 trials. These effect sizes were selected as our smallest effect size of interest based  | Primary tasksIf t-test is significant at p < 0.05, and indicates a difference in performance between the PL dual-task condition vs its standalone version and the VSSP dual-task condition, we will look at the means to conclude if the PL is involved in the processing of Arabic digits. If t-test is significant at p < 0.05, and indicates that performance is different in the PL dual-task condition vs the standalone condition  | If we observe PL involvement in either the primary or secondary task analysis, we will conclude that participants use verbal labels in the processing of Arabic digits. If we observe VSSP involvement in either the primary or secondary task analysis, we will conclude that  |

*Table C continued*

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Question** | **Hypothesis** | **Sampling plan** | **Analysis plan** | **Rationale** | **Interpretation** | **Theory** |
|  | compared to the PL standalone condition. |  | **Symbolic comparison (RT)**Paired sampled t-tests comparing: 1. Symbolic primary task with PL dual-task vs symbolic primary task standalone
2. Symbolic primary task with PL dual-task vs symbolic primary task with VSSP dual-task
3. Symbolic primary task with VSSP dual-task vs symbolic primary task standalone

Secondary tasksPaired samples t-test (standalone PL secondary task vs PL secondary task during dual-task accuracy; standalone VSSP secondary task vs VSSP secondary task during dual-task accuracy) | on adults’ performance of the primary and secondary tasks in prior research (Lyons et al., 2012; Maloney et al., 2019). | but the t-test indicates that there is no significant difference between the PL dual-task condition and the VSSP dual-task condition then we will conclude that there is an additional WM load from completing two tasks simultaneously, however we cannot be specific about the component of WM involved.Secondary tasksIf performance in the secondary task is significantly different between the PL dual-task condition and the PL standalone condition, we will look at the means to infer if the PL is involved in the processing of Arabic digits.If performance in the secondary task is significantly different between the VSSP dual-task condition and the VSSP standalone condition, we will look at the means to infer if the PL is involved in the processing of Arabic digits. | participants use visual strategies in the processing of Arabic digits.If we involve both VSSP and PL involvement, we will conclude that WM is required in processing Arabic digits, however we cannot be specific about which component.If we observe no WM involvement (either PL or VSSP), we will conclude that Arabic digits may be processed automatically. |

*Table C continued*

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Question** | **Hypothesis** | **Sampling plan** | **Analysis plan** | **Rationale** | **Interpretation** | **Theory** |
| **RQ1b** Is the processing of non-symbolic representations automatic or does it require the involvement of WMcomponents?  | **RQ1a.** Processing of non-symbolic representations will require the involvement of the VSSP but not the PL. This means we will expect to see a difference in performance in the **non-symbolic primary task** in the VSSP dual-task condition compared to the PL dual-task condition and standalone condition, OR a difference between the **VSSP secondary task** in the dual-task condition when compared to the VSSP standalone condition | Based on the effect size calculations described above, the smallest effect size of interest for this RQ is 0.35.Using a power level of 90%, an alpha level of 0.05 and a minimum effect size of Cohen’s d = 0.35, the minimum sample required for this RQ is 72. | Primary tasks**Non-symbolic comparison (accuracy)**Paired sample t-test comparing: 1. Non-symbolic primary task with VSSP dual-task vs non-symbolic primary task standalone
2. Non-symbolic primary task with VSSP dual-task vs non-symbolic primary task with PL dual-task
3. Non-symbolic primary task with PL dual-task vs non-symbolic primary task standalone

**Non-symbolic comparison (RT)**Paired sample t-test comparing: 1. Non-symbolic primary task with VSSP dual-task vs non-symbolic primary task standalone
2. Non-symbolic primary task with VSSP dual-task vs non-symbolic primary task with PL dual-task
3. Non-symbolic primary task with PL dual-task vs non-symbolic primary task standalone
 |  | Primary tasksIf t-test is significant at p < 0.05, and indicates that performance is different between the VSSP dual-task condition, the standalone condition and the PL dual-task condition, we will look at the means to determine if the VSSP but not the PL is involved in the processing of non-symbolic quantities. If t-test is significant at p < 0.05, and indicates that performance is different between the VSSP dual-task condition and the standalone condition but the t-test indicates that there is no significant difference between the VSSP dual-task condition and the PL dual-task condition, then we will conclude that thereis an additional WM load from completing two tasks simultaneously, however we cannot be specific about the component of WM involved. | If we observe PL involvement in either the primary or secondary task analysis we will conclude that participants use verbal labels in the processing of Arabic digits. If we observe VSSP involvement in either the primary or secondary task analysis, we will conclude that participants use visual strategies in the processing of non-symbolic quantities.If we observe both VSSP and PL involvement, we will conclude that WM is required in processing non-symbolic quantities, however we cannot be specific about which component.If we observe no WM involvement (either PL or VSSP), we will conclude that  |

*Table C continued*

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Question** | **Hypothesis** | **Sampling plan** | **Analysis plan** | **Rationale** | **Interpretation** | **Theory** |
|  |  |  | Secondary tasksPaired samples t-test (standalone VSSP secondary task vs VSSP secondary task during dual-task accuracy, standalone PL secondary task vs PL secondary task during dual-task accuracy) |  | Secondary tasksIf performance in the secondary task is significantly different in the VSSP dual-task condition than in the VSSP standalone condition, we will look at the means to determine if the VSSP is involved in the processing of non-symbolic quantities.If performance in the secondary task is significantly different between the PL dual-task condition and the PL standalone condition, we will determine if the PL is involved in the processing of non-symbolic quantities. | non-symbolic quantities may be processed automatically. |
| **RQ2** Can adults translate between Arabic and non-symbolic representations automatically or does this require access to verbal representations? | Translation between Arabic and non-symbolic representations will require the involvement of the PL. This means we will expect to see a difference in performance in the **cross-modal primary task** in the PL dual-task condition compared to the VSSP dual-task condition and standalone condition, OR a difference between the PL secondary task in the dual-task condition when  | Based on the effect size calculations described above, the smallest effect size of interest for this RQ is 0.33.Using a power level of 90%, an alpha level of 0.05 and a minimum effect size of Cohen’s d = 0.33, the maximum sample required for this RQ is 81. | Primary tasks**Cross-modal comparison (accuracy)**Paired sample t-test comparing: 1. Cross-modal comparison with PL dual-task vs cross-modal standalone
2. Cross-modal comparison with PL dual-task vs cross-modal with VSSP dual-task
3. Cross-modal comparison with VSSP dual-task vs cross-
 | The calculated effect sizes represent either: a reduction in primary accuracy of 5 trials, or a reduction in secondary accuracy of 2 trials. | Primary tasksIf t-test is significant at p < 0.05, and indicates that performance is different between the PL dual-task condition, the standalone condition and the VSSP dual-task condition, we will determine if the PL (but not the VSSP) is involved in the translation between Arabic digits and non-symbolic quantities. If t-test is significant at p < 0.05, and indicates that performance is different  | If we observe PL involvement in either the primary or secondary task analysis we will conclude that participants use verbal labels in translating between Arabic digits and non-symbolic quantities. If we observe VSSP involvement in either the primary or  |

*Table C continued*

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Question** | **Hypothesis** | **Sampling plan** | **Analysis plan** | **Rationale** | **Interpretation** | **Theory** |
|  | compared to the PL standalone condition |  | modal comparison standalone**Cross-modal comparison (RT)**Paired sample t-test comparing: 1. Cross-modal comparison with PL dual-task vs cross-modal standalone
2. Cross-modal comparison with PL dual-task vs cross-modal with VSSP dual-task
3. Cross-modal comparison with VSSP dual-task vs cross-modal comparison standalone

Secondary tasksPaired samples t-test (standalone PL vs dual-task PL accuracy; standalone VSSP vs dual-task VSSP accuracy) |  | between the PL dual-task condition and the standalone condition but the t-test indicates that there is no significant difference between the PL dual-task condition and the VSSP dual-task condition then we will conclude that there is an additional WM load from completing two tasks simultaneously, however we cannot be specific about the component of WM involved.Secondary tasksIf performance in the secondary task is significantly different between the PL dual-task condition than in the PL standalone condition, we will conclude that the PL is involved in translating between Arabic digits and non-symbolic quantitiesIf performance in the secondary task is significantly different between the VSSP dual-task condition and the VSSP standalone condition, we will look at the means to determine if the VSSP is involved in translating between Arabic digits and non-symbolic quantities | secondary task analysis, we will conclude that participants use visual strategies in translating between Arabic digits and non-symbolic quantities. If we observe both VSSP and PL involvement, we will conclude that WM is required in translating between Arabic digits and non-symbolic quantities, however we cannot be specific about which component.If we observe no WM involvement (either PL or VSSP), we will conclude translating between Arabic digits and non-symbolic quantities may be automatic. |

*Table C continued*

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Question** | **Hypothesis** | **Sampling plan** | **Analysis plan** | **Rationale** | **Interpretation** | **Theory** |
| **RQ 3a.** Does processing of Arabic digits differ for small and large quantities? | In the **symbolic comparison** condition, we expect no difference in the processing of small and large quantities. For both small and large quantities, we expect to see PL involvement. This means we will expect to see a difference in the **symbolic primary task** performance between the PL dual-task condition and the standalone condition, for both small and large quantities.  | Based on the effect size calculations described above, the smallest effect size of interest for this RQ is 0.5.Using a power level of 90%, an alpha level of 0.05 and a minimum effect size of Cohen’s d = 0.5, the maximum sample required for this RQ is 36. | For small quantitiesPaired samples t-test (standalone symbolic comparison vs symbolic comparison with PL dual-task)For large quantitiesPaired samples t-test (standalone symbolic comparison vs symbolic comparison with PL dual-task) | The calculated effect sizes represent either: a reduction in primary accuracy of 5 trials or a reduction in secondary accuracy of 2 trials. | If t-tests for both small and large quantities are significant (p < .05) we will conclude that processing both small and large Arabic digits involves the PL. | If t-tests for either small or large quantities are not significant, we will conclude that processing of these quantities may be automatic. |
| **RQ 3b.** Does processing of non-symbolic representations differ for small and large quantities? | In the **non-symbolic comparison** condition, we expect that small quantities will be processed automatically, whilst large quantities will involve the VSSP. This means we will expect to see a difference in the **non-symbolic primary task** performance between the VSSP dual-task condition and the standalone condition for large quantities, but no decrease in performance for small quantities. | Using a power level of 90%, an alpha level of 0.05 and a minimum effect size of Cohen’s d = 0.35, the maximum sample required for this RQ is 72. | For small quantitiesPaired samples t-test (standalone non-symbolic comparison vs non-symbolic comparison with VSSP dual-task)For large quantitiesPaired samples t-test (standalone non-symbolic comparison vs non-symbolic comparison with VSSP dual-task) | The calculated effect sizes represent either: a reduction in primary accuracy of 5 trials, or a reduction in secondary accuracy of 2 trials. | If t-tests for both small and large quantities are significant (p < .05) we will conclude that processing both small and large Arabic digits involves the VSSP. | If t-tests for either small or large quantities are not significant, we will conclude that processing of these quantities may be automatic. |

*Table C continued*

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Question** | **Hypothesis** | **Sampling plan** | **Analysis plan** | **Rationale** | **Interpretation** | **Theory** |
| **RQ 3c.** Does translation between Arabic and non-symbolic representations differ for small and large quantities? | In the **cross-modal comparison** condition, we expect no difference in the processing of small and large quantities. For both small and large quantities, we expect to see PL involvement. This means we will expect to see a difference in **cross-modal primary task** performance between the PL dual-task condition and the standalone condition, for both small and large quantities.  | Using a power level of 90%, an alpha level of 0.05 and a minimum effect size of Cohen’s d = 0.33, the maximum sample required for this RQ is 81. | For small quantitiesPaired samples t-test (standalone cross-modal comparison vs cross-modal comparison with PL dual-task)For large quantitiesPaired samples t-test (standalone cross-modal comparison vs cross-modal comparison with PL dual-task) | The calculated effect sizes represent either: a reduction in primary accuracy of 5 trials or a reduction in secondary accuracy of 2 trials. | If t-tests for both small and large quantities are significant (p < .05) we will conclude that translating between Arabic digits and non-symbolic quantities involves the PL for both small and large quantities. | If t-tests for either small or large quantities are not significant, we will conclude that translating between these quantities is automatic. |

**Appendix D**

***Exploratory Analyses***

RQ3 regarded the possible differential impact of interference for small and large quantities and numbers. In this case, our preregistered analyses were dependent on our predictions regarding which WM component would be involved with each type of primary task (RQs 1-2). So, for symbolic and cross-modal comparison, we had hypothesised PL involvement and therefore only preregistered investigating the impact of interference separately for small and large-number trials for PL interference. Similarly, for non-symbolic comparison, we had only hypothesised VSSP WM involvement. Our results, however, were more nuanced. To present the complete picture of the effects of the interference conditions, in Figure 5, we also report the results of paired sample t-tests for comparisons of with and without interference separately for small and large number-trials in the interference conditions that were not originally expected (i.e. in the case of symbolic and cross-modal comparison for PL, and in the case of non-symbolic comparison for VSSP). These are presented with grey arrows in Figure 5. These results should be interpreted with caution given their exploratory nature.

 Further inspection of the overall results in Figure 5, especially in the case of the symbolic comparison primary task, one may wonder whether our findings were affected by a speed-accuracy trade-off. Specifically, we found that under both interference conditions in the symbolic task participants were more accurate but slower compared to the standalone condition. Initially, we had hypothesised that significant increases in RT would be reflective of interference effects, but in this case, rather than slowing because of the interference, this RT effect may have resulted from improved attention to the primary task under interference and being more accurate. To test this assumption, we ran the respective correlations (Table D3). We found a significant but small correlation between accuracy and RT only in the case of PL interference but this correlation did not significantly differ from the respective symbolic standalone and symbolic VSSP speed-accuracy r values, indicating that there was no significant change in a speed-accuracy trade-off across these conditions (Standalone vs. PL interference: z = - 0.78, *p* = .43; Standalone vs. VSSP interference: z = 0.49, *p =* 0.62; PL interference vs. VSSP interference: z = -1.76, *p* = .08).

**Table D3.**

*Exploring potential speed-accuracy trade-offs across the symbolic conditions*

|  |  |  |
| --- | --- | --- |
| **Correlation Accuracy-RT** | **r** | **p-value** |
| Symbolic standalone | 0.20 | .07 |
| Symbolic PL | 0.30 | .009 |
| Symbolic VSSP | 0.13 | .26 |

Lastly, it is well known that adults perform best in symbolic, then non-symbolic and finally cross-modal comparison tasks (Lyons et al., 2012). We checked and confirmed that this was the case with our accuracy data too, *F* (2, 240) = 26.47, *p* < .001 (*MSymbolic standalone* = 0.96, *SD* = 0.03; *MNon-symbolic standalone* = 0.93, *SD* = 0.06; *MCross-modal standalone* = 0.91, *SD* = 0.06; post-hoc comparisons: Symbolic vs. Non-symbolic: t = 4.01, *p* < .001; Non-symbolic vs. Cross-modal: t = 3.25, *p* = .001; Symbolic vs. Cross-modal: t = 7.26, *p* < .001).

1. This minor change from the Stage 1 manuscript was approved by PCI RR on 16/03/2023. [↑](#footnote-ref-2)