RUNNING HEAD: Processing and translating of numerical representations

Stage 1 Registered Report: Can adults automatically process and translate between numerical representations?

Iro Xenidou-Dervou [a]*, Caroline Appleton [a], Camilla Gilmore [a]

[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. IXD will write the Stage 2 manuscript. XXX[1] will recruit participants and collect the data. CA, IXD and CG will perform all analyses.

Conflict of interest statement:

Authors declare no conflict of interest

Data availability statement:

Raw data will be available on the Open Science Framework upon submission of the Stage 2 manuscript and will be made publicly available following publication of the Stage 2 manuscript.

[1] To be confirmed

Can adults automatically process and translate between numerical representations?

Abstract

Arithmetic, and the ability to use numbers, is an important skill in adult life. Numbers can be represented in three ways: through number words, in a visual Arabic number form and non-symbolically. Much research attention has focussed 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 resources. In this registered report, we propose a study using the dual-task paradigm to answer this question.

We will administer dot comparison, digit comparison and cross-modal comparison tasks as primary tasks, which will be conducted in standalone and dual-task (phonological and visuospatial) conditions. Examining performance across all three primary tasks allows us to draw conclusions about the specific nature of numerical representations both when processing and translating different representations. 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.

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 life 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 the Approximate Number System (ANS; e.g., Halberda & Feigenson, 2008). The ANS is assumed to provide estimates of the quantity that a given set of nonsymbolic 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., 2013)

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 representations. As described above, numbers can be represented with number words (e.g., "three"), through Arabic symbols (3), or, as termed by Dehaene, 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 symbol and non-symbolic quantity) than when completing dot and digit comparisons (processing of non-symbolic quantities/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 are 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 inaccurate 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 Arabic 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 the physically larger, the trials are congruent and reactions times are lower. However, where trials are incongruent, reaction times are higher (Reike & Schwarz, 2017).

From this we can infer 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 on 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 automaticity of 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 tasks 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 processing information in specific modalities (Baddeley, 2010), and the central executive (CE), which controls the two 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, van der Schoot, et al., 2015). 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 measures of the ANS, 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. However, correlational studies cannot tell us whether WM resources are required for processing (i.e., comparing or manipulating numerical representations within a particular code: verbal, Arabic, 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 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 stand-alone conditions, i.e. without a dual-task load (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 socalled 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 aims to investigate the processing of Arabic symbols and non-symbolic quantities, and the role of verbal representations in translating between Arabic 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 can expect to see no WM involvement.

To examine the processing of numerical representations, we will administer dot comparison, digit comparison and cross-modal comparison tasks as primary tasks, which will be conducted in standalone and dual-task (phonological and visuospatial) conditions. This allows 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 allows us to draw conclusions about the specific nature of numerical representations both when processing and translating different representations. The use of three comparison tasks allows us to draw conclusions about the nature of each representation and ensures that any WM involvement is 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 allows us to answer further questions about the nature of representations in each modality. We expect 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 aim 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 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 will therefore examine the differences in how small (1-4) and larger (5-9) quantities are processed and translated. We choose these quantities, rather than quantities greater than 10, as whilst the non-symbolic quantities are inaccurate, it is still possible for adults to attach Arabic symbols to these quantities. We expect that quantities in the small range will involve more phonological processes than those in the large range, because small non-symbolic representations are 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 aim to address the following research questions:

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

Resources (analysis plans, stimuli, experiment scripts) for the following experiment can be found here: https://bit.ly/3lFeWll

Method

Participants.

Adult participants (age 18-65) will be recruited via university email and social media. Research has shown that there is relatively little change in adults' WM performance within this age-range (Alloway & Alloway, 2013). Participants must have normal or corrected-to-normal vision and hearing and speak English as their first language. Ethical approval has been granted by Loughborough University Ethics Committee and participants will be reimbursed for their time.

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 will be 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 is the sample size we will recruit for this experiment. For RT, this would allow us to detect differences of 50ms for the symbolic and nonsymbolic comparison conditions and a difference of 80ms for the cross-modal comparison condition. For secondary tasks, there will be a total of 20 trials. A decrease in 1 sequence length (e.g., remembering 6 items versus remembering 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 expect 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 can also inevitably be considered arbitrary since no study has previously examined these effects, therefore nonsignificant results will 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 a document on the OSF page.

Materials

Primary tasks.

Numerical comparison tasks. Participants will complete symbolic, non-symbolic and cross-modal comparison tasks. The quantities used in each task will be the same. Small numbers comprise 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) will be used, with the exception of pairs with a ratio of 0.25, which will be removed in order to equate difficulty across the small and large sets. 11 will be added to the large set, to ensure that participants do not always select 9 as the larger quantity, however these trials will be 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 will be counterbalanced.

Quantities will be presented on the screen, and participants are instructed to select the larger quantity and respond using the keyboard (“z” if the left quantity is larger, “m” if the right quantity if larger). Quantities will appear on the screen for 1000ms, to prevent counting, however participants can 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 are congruent with quantity (i.e. the array with the higher quantity of dots also has a larger diameter, greater density etc.), and in half of the trials visual parameters are incongruent with dot quantity (Wang et al., 2020).

Secondary tasks.

Phonological. A reverse letter span task will be used as a secondary task to load the PL component of WM. The sequence of events will be as follows: 1) Participants will be presented with a randomised sequence of letters (1 second per letter, presented orally through the computer) and told to remember the sequence. Each letter can 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 will then be 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 requires participants to use their WM to process the information, as opposed to simply maintaining the letters in short-term memory.

The span will range 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 will be used for each span length, resulting in a total of 20 trials. For the secondary task, we will record accuracy of recall.

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

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

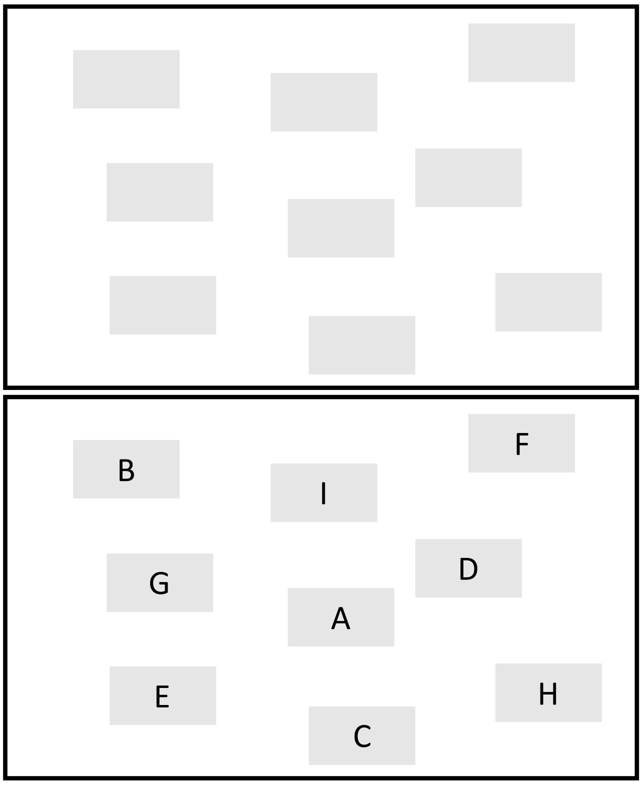


Figure 1. *VSSP secondary task. Above image shows the block presentation at the start of each trial. Below image shows the blocks at the end of the trial, with letters added to allow the participant to recall the sequence verbally.*

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

Procedure

All participants will complete all conditions. The order of primary tasks will be randomised within PsychoPy, and the order of secondary tasks will be counterbalanced within participants. This means that participants will complete each primary task in standalone and dual-task conditions, before moving on to the next primary task. Participants will also complete both secondary tasks as standalones. An example of the procedure, demonstrated for the non-symbolic comparison condition, is shown in Figure 2.

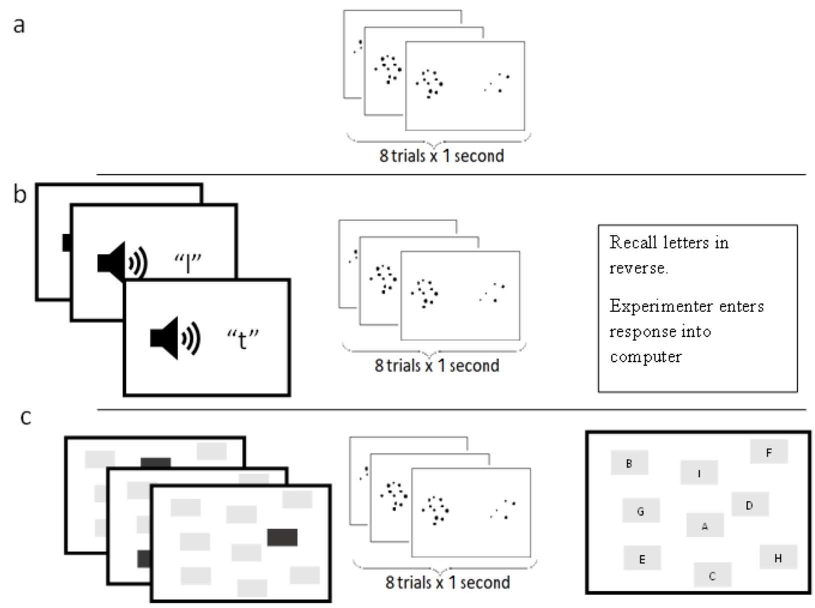


Figure 1. Example of conditions for non-symbolic comparison. Condition (a) shows standalone primary, (b) shows phonological dual-task and (c) shows visuospatial dual-task

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 will examine performance in both tasks, in comparison to the standalone conditions.

Research questions 1 and 2 will be 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 conditions. These will be performed separately for the different primary tasks (non-symbolic comparison, symbolic comparison, cross-modal comparison) and secondary tasks (PL, VSSP). Initial analyses will be conducted for small and large trials combined before further analyses consider small and large trials separately, to answer Research Question 3. A summary of the analysis plan can be found in Appendix A, with full details available on OSF (https://bit.ly/3lFeWll).

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

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

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

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

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

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

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

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

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

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

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

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

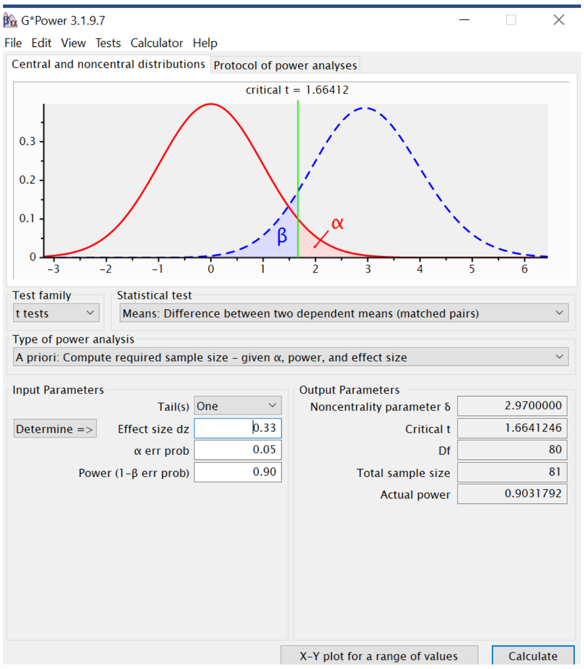
Appendix A

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

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 will be conducted in JASP.

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.

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	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	<u>Primary tasks</u> **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 on adults' performance of	<u>Primary tasks</u> If 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 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

	standalone condition, OR a difference between the **PL secondary task** in the dual-task condition when compared to the PL standalone condition.	required for this RQ is 36.	**Symbolic comparison (RT)** Paired sampled t-tests comparing: 4) Symbolic primary task with PL dual-task vs symbolic primary task standalone 5) Symbolic primary task with PL dual-task vs symbolic primary task with VSSP dual-task 6) Symbolic primary task with VSSP dual-task vs symbolic primary task standalone <u>Secondary tasks</u> Paired 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)	the primary and secondary tasks in prior research (Lyons et al., 2012; Maloney et al., 2019).	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 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. <u>Secondary tasks</u> If 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	conclude that 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.

Deleted: are

					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.	
RQ1b Is the processing of non-symbolic representations automatic or does it require the involvement of WM components?	**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	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.	<u>Primary tasks</u> **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		<u>Primary tasks</u> If 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 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

	between the **VSSP secondary task** in the dual-task condition when compared to the VSSP standalone condition		primary task standalone **Non-symbolic comparison (RT)** Paired sample t-test comparing: 4) Non-symbolic primary task with VSSP dual-task vs non-symbolic primary task standalone 5) Non-symbolic primary task with VSSP dual-task vs non-symbolic primary task with PL dual-task 6) Non-symbolic primary task with PL dual-task vs non-symbolic primary task standalone <u>Secondary tasks</u> Paired 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)		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 there is an additional WM load from completing two tasks simultaneously, however we cannot be specific about the component of WM involved. <u>Secondary tasks</u> If 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	visual strategies in the processing of non-symbolic quantities. If we involve 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 non-symbolic quantities may be processed automatically.

Deleted: are

					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.	
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	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-modal comparison standalone	The calculated effect sizes represent either: a reduction in primary accuracy of 5 trials, or a reduction in secondary accuracy of 2 trials.	Primary tasks If 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 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 secondary task analysis, we will

	between the **PL secondary task** in the dual-task condition when compared to the PL standalone condition		**Cross-modal comparison (RT)** Paired sample t-test comparing: 4) Cross-modal comparison with PL dual-task vs cross-modal standalone 5) Cross-modal comparison with PL dual-task vs cross-modal with VSSP dual-task 6) Cross-modal comparison with VSSP dual-task vs cross-modal comparison standalone Secondary tasks Paired samples t-test (standalone PL vs dual-task PL accuracy; standalone VSSP vs dual-task VSSP accuracy)		If t-test is significant at $p < 0.05$, and indicates that performance is different 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 tasks If 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	conclude that participants use visual strategies in translating between Arabic digits and non-symbolic quantities. If we involve 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.

Deleted: is

					translating between Arabic digits and non-symbolic quantities 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 determine if the VSSP is involved in translating between Arabic digits and non-symbolic quantities	
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	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	For small quantities Paired samples t-test (standalone symbolic comparison vs symbolic comparison with PL dual-task) For large quantities Paired 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.

Deleted: is

	between the PL dual-task condition and the standalone condition, for both small and large quantities.	maximum sample required for this RQ is 36.				
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 quantities Paired samples t-test (standalone non-symbolic comparison vs non-symbolic comparison with VSSP dual-task) For large quantities Paired 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.
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	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	For small quantities Paired samples t-test (standalone cross-modal comparison vs cross-modal comparison with PL dual-task) For large quantities	The calculated effect sizes represent either: a reduction in primary accuracy of 5 trials or a reduction in secondary	If t-tests for both small and large quantities are significant ($p < .05$) we will conclude that translating between Arabic digits and non-symbolic	If t-tests for either small or large quantities are not significant, we will conclude that translating between these quantities may be automatic.

Deleted: is

Deleted: is

	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.	required for this RQ is 81.	Paired samples t-test (standalone cross-modal comparison vs cross-modal comparison with PL dual-task)	accuracy of 2 trials.	quantities involves the PL for both small and large quantities.	