**The origin of symbolic numerical knowledge in early development –**

**An fNIRS Registered Report**

Elizaveta Ivanova1,\*, Marc Joanisse2, Daniel Ansari2, Mojtaba Soltanlou1,3

1 School of Psychology, University of Surrey, UK

2 Brain and Mind Institute and Department of Psychology, Western University, Canada

3 Department of Child Education, Faculty of Education, University of Johannesburg, South Africa

\* Corresponding:

Elizaveta Ivanova

elizaveta.ivanova.w@gmail.com

School of Psychology, University of Surrey, UK

Running header: The origin of symbolic numerical knowledge in preschool children

Keywords: preschool children, numerical cognition, cardinality principle knowledge, functional near-infrared spectroscopy (fNIRS), functional connectivity

**Abstract**

The ability to understand that each number word in the count sequence refers to a specific set of items (e.g., ‘two’ means ‘two things’) is a milestone in cognitive development. When children reach this milestone, they are said to understand the cardinality principle: the last word in the count sequence refers to the total number of items in the set. Acquisition of the cardinality principle is the most crucial step in the development of symbolic numerical knowledge. While research in cognitive development has revealed much about how and when children learn the cardinality principle, we know very little about the neural changes that are associated with this major conceptual development. This understanding will provide evidence for developing theories about the origin of human symbolic knowledge and individual differences in early development, which might help to identify early risk factors for later math learning difficulties. In this study, we will investigate the bilateral parietal responses using functional near-infrared spectroscopy (fNIRS) during the numerical adaptation task in 2 years 9 months and 4 years 7 months old children. While one group has already acquired cardinality principle knowledge (understanding that each number word associates with a corresponding magnitude), the other group has not yet acquired it. We will apply univariate analysis along with functional connectivity to compare brain responses in the two groups. We hypothesize different brain activation patterns and connectivity between the two groups: children who demonstrate knowledge of the cardinality principle will reveal higher left parietal activation related to the semantic processing of number words and higher bilateral parietal connectivity related to a link between number words to their discrete as compared to children who have not acquired this knowledge yet.

**Introduction**

The ability to understand that each number word in the count sequence refers to a specific set of items (e.g., ‘two’ means ‘two things’) is a milestone in cognitive development (e.g., Haman et al., 2023; Merkley & Ansari, 2016). When children understand that each word in their count sequence refers to a specific number of items in a set, they are said to understand the cardinality principle (Gallistel & Gelman, 1990; Le Corre & Carey, 2007; Sarnecka et al., 2015; Wynn, 1990). In other words, they understand that the last word in the count sequence represents the total number of items in the counted set (e.g., when they count all the chocolates in a bowl, they know the total number of chocolates is the last number they count). Acquisition of the cardinality principle is the most crucial step in the development of symbolic numerical knowledge, which predicts how well children go on to develop numerical and mathematical skills (Geary et al., 2018). While research in cognitive development has revealed much about how and when children learn the cardinality principle, we know very little about the neural changes that accompany this major conceptual development. This insight would provide a better understanding of individual differences in knowledge acquisition in early development that feeds theories about the origin of human symbolic knowledge.

The development of the cardinality principle occurs in a known trajectory. Before acquiring cardinality principle knowledge, children can verbally count without knowing what number words mean and what they refer to. Instead, they repeat the memorized count sequence they have learned from adults and older children. For example, young children may count as they walk up a flight of stairs but are not yet able to tell how many steps they have climbed. Children gradually begin to enumerate sets by understanding what the purpose of counting is (e.g., to understand how many chocolates they are allowed to pick from the bowl). The gradual emergence of the cardinality principle knowledge begins when children first understand that sets of one and two refer to the quantities of ‘one’ and ‘two’. Children become so-called ‘one’ and ‘two’ knowers approximately between the ages of 2-3. They then move on to being ‘three’, then ‘four’ knowers, and so on every few months (Sarnecka et al., 2015). By extending this knowledge to numbers above ‘four’, children become cardinal principle knowers. After the acquisition of the cardinality principle, children will know that each number word in their counting refers to a specific quantity of items. Typically, children are thought to have developed a fully-fledged understanding of the cardinality principle by about age 5, however, some children may not be cardinality principle knowers at this age.

Cardinality principle knowledge in children is most commonly evaluated by the give-a-number task (Wynn, 1990). In this task, children are asked to give the experimenter an exact number of objects from a pile of toys (Sarnecka et al., 2015). For instance, the experimenter gives a basket of small toys to the child and asks them “Could you give me *n* toys?” and the child has to count the toys one by one while giving them to the experimenter until they reach *n* toys. The give-a-number task determines number-knower levels including one-knower, two-knower, three-knower, four-knower, and cardinality principle knowers (Sarnecka et al., 2015). Children who can correctly give the experimenter 1-4 objects, but not more, are referred to as subset-knowers because they are not able to use counting to determine the numerical quantity of the sets above four. These children are not considered to have full-fledged knowledge of the cardinality principle. Subset-knowers might verbally count up to a large number, but in the give-a-number task, they are only able to give the correct number of items for sets up to ‘four’. In contrast, cardinality principle knowers (CP-knowers) are those who have already achieved cardinality principle knowledge and can give the correct number for every number in their count list. Put differently, CP-knowers understand that counting can be used to determine the numerical quantity of any set.

In contrast to the wealth of behavioural studies into the developmental processes associated with the acquisition of cardinality principle knowledge, we know very little about the neurocognitive mechanisms underpinning cardinality knowledge (Hyde, 2021). In a relevant ERP study, Pinhas et al. (2014) observed greater negativity over centroparietal sites and greater positivity over bilateral parietal sites during an audio-visual matching task in CP-knowers as compared to subset-knowers. They reported an engagement of the parietal region in linking number words to sets of items in CP-knowers (Pinhas et al., 2014). This engagement of the bilateral parietal regions during number processing has been reported in preschool children who acquired cardinal principle knowledge (Bugden, et al., 2021; Cantlon et al., 2006; Cantlon et al., 2011). That being said, the right and left parietal regions provide support for the numeracy development unilaterally as well. For the right parietal cortex, neuroimaging studies in early childhood have consistently shown response during non-symbolic processing (changes in the number of presented dots; Cantlon et al., 2006; Edwards et al., 2016; Hyde & Spelke, 2011; Hyde et al., 2010; Izard et al., 2008). For the left parietal cortex, neuroimaging studies in school children and adults revealed the left parietal response during processing number words (e.g., Vogel et al., 2017). While these studies provide valuable insights about parietal engagement in number processing, no study has investigated the transition from subset knowledge to cardinality knowledge.

In addition to the parietal cortex, research has frequently shown prefrontal engagement during number processing (Arsalidou et al., 2011, 2018). In infancy, frontal activation has been observed in response to error detection or violation of expectation during non-symbolic numerical tasks (e.g., adding one puppet to another resulted in one puppet; Berger et al., 2011, 2019). At preschool age, children who solved a number conservation task (i.e., the number of items is identical no matter how far away they are stretched on the screen from each other) showed higher activation in the frontal region compared to children who were not able to solve the problem (Simon et al., 2015). Similarly, 5-6-year-old children showed frontal activation when completing a numerical Stroop task (Ben-Shalom et al., 2013). These studies suggest that bilateral frontal regions are involved in the development of numerical understanding in the early years of life. However, both for the frontal and parietal regions, most of the evidence does not tie to a significant behavioural change such as the transition from subset-knower to CP-knower. More importantly, most neuroimaging studies of cardinality are with older children whom we expect have already progressed to the cardinality principle stage. The litmus test would be whether the fundamental conceptual development that is represented by the acquisition of the cardinality principle is associated with functional organizations in the frontoparietal network of mathematical thinking during the critical age of 3-4 years.

The development of the cardinality principle might be also associated with a functional coupling between the bilateral parietal and frontal regions. The bilateral frontoparietal connectivity represents a connection between number words in the left parietal cortex, more intuitive representations of quantity in the right parietal cortex and mapping between quantities and symbols in the bilateral frontal cortex (Hyde, 2021; Sokolowski et al., 2017, Hubbard et al., 2008; Rivera et al., 2005, Cantlon et al., 2009). An fMRI study by Emerson and Cantlon (2011) showed that frontoparietal connectivity is correlated with both basic number matching ability and scores on the standardized test of mathematical ability in 4-11-year-old children. This finding suggests the functional connectivity of the network is math-specific, and can potentially be a predictor of mathematics achievement. Further study nuanced the developmental shifts in connectivity of the frontoparietal network: a resting-state fMRI study (Zhang et al., 2019) showed that the frontoparietal numerical network has a strong association with counting at 4 years old, while gradually shifting to parietal regions over time, as early as 6 years old. Taken together, the frontoparietal network is suggested to vastly affect numerical processing, accommodating both its symbolic and non-symbolic aspects. It can also be suggested that the role of the frontal region in the frontoparietal connectivity is strongly pronounced in the first years and is being shifted to parietal connectivity every year, as the numeracy skill develops (Hyde, 2021). Therefore, the cardinality principle knowledge might be associated with greater functional connectivity between frontoparietal regions in 3-4-year-old children.

The current study aims to advance our understanding of the brain mechanisms underpinning children’s *acquisition* of cardinality principle knowledge. Using the give-a-number task, we will first screen preschool children between the age of 2 years 9 months and 4 years 7 months old to distinguish CP-knowers and subset-knowers. Then they will perform an auditory number word adaptation task (adapted from Vogel et al., 2017) while their brain responses are recorded. In this task, children will hear the number word ‘two’ presented repeatedly, occasionally interspersed with deviant auditory number words (i.e., ‘four’ or ‘eight’) or a deviant auditory non-number word (i.e., ‘rin’ which is a phonotactic legal anagram of the number word). The task does not demand active response production that mitigates the influence of response selection (Vogel et al., 2015) and makes it suitable for young children (Kovacs & Schweinberger, 2016). We expect different brain responses in CP-knowers and subset-knowers to the deviant auditory number words but not to the deviant auditory non-number word. CP-knowers process the conceptual difference between repeated number words ‘two’ and deviant auditory number words ‘four’ and ‘eight’ that points to the association between the semantics and quantity of those numbers, which is not the case for subset-knowers. In other words, the number words ‘four’ and ‘eight’ represent a set of items in the brain in CP-knowers but not in subset-knowers. This conceptual difference is beyond the perceptual difference that ‘two’ sounds different from ‘four’ and ‘eight’ and all of them belong to their count sequence, which is true for both CP-knowers and subset-knowers. Brain activation will be measured using functional near-infrared spectroscopy (fNIRS), which is non-invasive, portable, low sensitive to movement, and low cost (Barreto & Soltanlou, 2022;Pinti et al., 2018; Scholkmann et al., 2014; Soltanlou, Sitnikova, et al., 2018). Those characteristics allow recording brain responses in natural settings like nurseries and, make it one of the most promising neuroimaging techniques in developmental science (e.g., Aslin et al., 2015; Soltanlou et al., 2017; Soltanlou, Artemenko, et al., 2018; Soltanlou et al., 2022; Whiteman et al., 2017).

We hypothesize that CP-knowers will exhibit higher left parietal activation, defined by increased HbO/decreased HHb, relative to subset-knowers because they are more advanced in their conceptual knowledge about the meaning of number words (Hypothesis 1). Moreover, CP-knowers will exhibit higher bilateral parietal functional connectivity, defined by functional co-activation between these two regions, relative to subset-knowers because they have built a link between number words and intuitive representations of quantity (Hypothesis 2). We summarised the direction of the hypotheses in Figure 1.

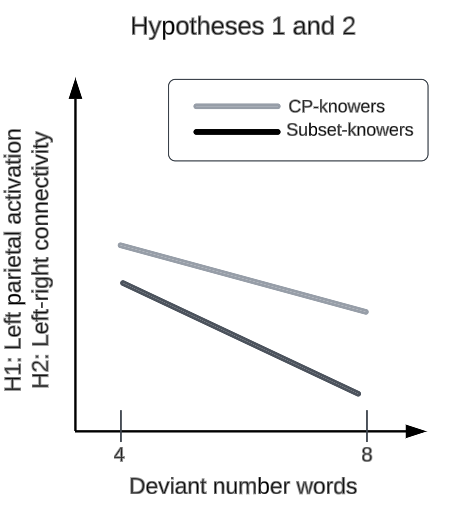
****

Figure 1: conceptual visualisation of the hypotheses for this study. According to Hypothesis 1, CP-knowers will exhibit higher left parietal activation, defined by increased HbO/decreased HHb, in both conditions relative to subset-knowers because they are more advanced in their conceptual knowledge about the meaning of number words. Similarly, according to Hypothesis 2, CP-knowers will exhibit higher bilateral parietal functional connectivity, defined by functional co-activation between these two regions. Please note that in both hypotheses, we expect smaller differences between CP-knowers and subset-knowers for the number word ‘four’ than for the number word ‘eight’. While both CP- and subset-knowers understand the number word ‘four’ and are expected to have a strong parietal response, CP-knowers are more advanced, so their parietal response will be stronger as compared to subset-knowers. However, a larger group difference for the number word ‘eight’ is expected because subset-knowers do not understand the number word ‘eight’ yet and are thus not expected to have a strong parietal response. CP-knowers, however, have a much better understanding of the number word ‘eight’, and are expected to have a strong parietal response. Please also note that the lines’ angle of inclination serves only illustrative purposes, namely, showing which group and which condition is suggested to have the highest activation (Hypothesis 1) or strongest connectivity (Hypothesis 2).

**Material and Methods**

**Bayes factor design analysis**

Since this is the first fNIRS study of numerical knowledge acquisition in preschool children, the sample size calculation is based on the most relevant study by Holloway et al. (2013). They had two groups of adult participants: bilingual in English and Chinese, and fluent in English but not Chinese. In one of their conditions, Chinese numbers were used. Therefore, while these numbers were familiar (i.e., semantically meaningful) for the bilingual participants, they all seemed to be unfamiliar for English monolingual participants. Holloway et al. (2013) observed significant group differences in the right IPS activation during this Chinese numerical adaptation task. We used the means and standard error values of bilingual (*M* = 0.406, *SE* = 0.101) and monolingual (*M* = 0.153, *SE* = 0.089) groups provided on p. 395-396 of Holloway et al. (2013) in the Chinese numerical adaptation condition. These values were used to calculate the standard deviation values (*SD* = 0.364 for bilingual, and *SD* = 0.321 for monolingual) and the Cohen’s *d* effect size of 0.73.

There are several methodological differences that should be taken into account when using their effect size for the sample size calculation. For instance, in the study by Holloway et al. (2013), two groups of adults performed a visual numerical adaptation task while brain activation was monitored by fMRI and the dependent variables were ratio-dependent changes (e.g., 6/8 vs. 6/12). In contrast, in our study, two groups of preschoolers will perform an auditory numerical adaptation task while brain activation will be monitored by fNIRS and the dependent variables will be the contrast of number words vs. non-number words (e.g., ‘four’ vs. ‘rin’). In addition, there is a considerable difference in signal-to-noise ratio for fMRI vs. fNIRS with different spatial and temporal resolutions, which can mean statistical power from an fMRI study may not be directly compared to that of fNIRS. Considering also the file drawer problem of unpublished studies, and publication bias towards significant findings (Cipora & Soltanlou, 2021; Fanelli, 2011), as well as the current lack of studies on the neuroimaging differences between subset-knowers and CP-knowers, it is sensibleto utilize a low to medium effect size of Cohen's *d* = 0.35.

We applied Bayes Factor Design Analysis (BFDA) using sequential Bayes factor with maximal n design, in which participants will be collected until either a sufficient BF is achieved, or a maximum number is reached (Schönbrodt & Wagenmakers, 2018). The benefit of using this design is that it ensures collecting sufficient sample while ensuring feasibility (Schönbrodt & Wagenmakers, 2018). Calculation of sample size involved running Monte Carlo simulations and relied on the web-based BFDA application (Stefan et al., 2019; see http://shinyapps.org/apps/BFDA/) Simulation results indicate that with informed prior distribution of effect sizes (t(mu = 0.35, r = 0.102, df = 3; Stefan et al., 2019), the expected sample size to reach at least BF10 of 6 is 46 participants per group. As the testing with 20 participants per group has been shown to highly likely cause misleading evidence (see again Stefan et al., 2019; Schönbrodt et al., 2017), we will start consecutive testing for the hypotheses once there are 25 participants per group and will stop once the testing reaches BF **>** 6 in favour of H1 or H0 or the maximum of 46 participants per group has been collected. Given that even 92 participants would be considered as a large sample for a neuroimaging study in preschoolers, if the BF for either of the hypotheses does not reach 6 (whether due to the effect’s extreme weakness or the common in the field deficit in neuroimaging data quality in preschoolers) even after recruiting maximum feasible sample size, we will consider this inconclusive result as an important message to the field.

**Participants**

We will recruit typically developing preschool children between 2 years 9 months and 4 years 7 months old through social media and nurseries in the UK. Note that we will continue recruitment until we reach the target. Due to the possibility of testing children in both English and Russian languages, we will invite the children who speak English or Russian as their primary language. In the case where the child speaks both languages equally fluently, the language of study will be defined based on the child’s preferences. Note that English and Russian languages have similar grammar structures regarding numerals, and native speakers of both languages have been found to develop numeracy skills within the same timeline (Sarnecka et al., 2007), hence, the language will not be a grouping factor in the analyses. Demographic information about children including age, biological sex, handedness, socioeconomic status, number of siblings, frequency of day-care attendance, and parents’ education will be collected.

Inclusion criteria consist of (i) age between 2 years 9 months and 4 years 7 months old at the time of measurement, (ii) no history or diagnosis of neuropsychological impairment including developmental disorder, (iii) no chronic disease such as diabetes mellitus, renal failure, and high blood pressure, (iv) no use of particular medication at the time of measurement, (v) normal or corrected-to-normal vision, and (vi) ability to speak and understand English and/or Russian. Exclusion criteria consist of (i) non-compliance with the wearing fNIRS cap (declining to wear the cap completely), (ii) non-completion of the Give-a-number task, (iii) extreme contamination of fNIRS data (rejection of all channels during preprocessing), (iv) experimental error during Give-a-Number task and fNIRS task. Parents will give written consent and receive a reimbursement, and similarly, children will give verbal consent and receive a toy for participation in the study. All procedures are in line with the latest ethics guidance of the Economic and Social Research Council (ESRC) and were approved by the University Ethics Committee of the University of Surrey.

**Questionnaires and behavioural measures**

*Socioeconomic status questionnaire (SES)*

The socioeconomic status will be measured using the Barratt Simplified Measure of Social Status (BSMSS) Questionnaire (Barratt, 2006), and it should be completed by Parents/Guardians. In this questionnaire, the socioeconomic status is represented by maternal and paternal years of education. Unlike income and occupation measurements, this construct is suggested to be both more stable and objective, as well as relatable to the quality of a child’s home environment (Duncan & Magnuson, 2012; Mayes & Lewis, 2012).

*Home-Maths Environment Questionnaire*

This questionnaire is to be conducted by the child’s parent/guardian to find out about math activities the child carries out at home. It will include a total of 20 items. The first 7 items measure the frequency of parent/guardian-child mathematical home activities. Questions regarding how often parent/guardian engages in different math-related activities with their child at home, such as cooking and playing games. Parents’ frequency of at-home math activity ranges on a 5-point Likert scale; never (0) to every day (4). The next 10 items measure parent/guardian expectations of their child’s mathematical skills. This includes questions about how important the parent/guardian felt it was for their child to be able to complete certain math-related tasks. This will be measured using a 4-point Likert scale ranging from not at all important (0) to very important (3). The final 3 items will measure parents/guardians’ attitudes towards mathematics using a 5-point Likert scale, ranging from completely disagree (0) to completely agree (4). For each participant, an average score will be calculated for each subscale. With a Cronbach’s alpha of 0.79, 0.89 and 0.74 for the activities, expectations and attitudes items respectively, this questionnaire has sufficient reliability. This questionnaire is based on De Keyser et al., 2020, with the work adapted from LeFevre et al., 2009, Kleemans et al., 2012 and Rathé et al., 2020.

*Non-verbal IQ measurement*

To assess the non-verbal IQ in both groups of children, the Matrices subtest of the British Ability Scales III (BAS III) will be administered (Elliot & Smith, 2011). It is designed to assess the non-verbal IQ of different abilities including visual-perceptual organization, nonverbal reasoning, and cognitive flexibility. In the task, children will be shown incomplete matrices of abstract figures. Children will need to complete the pattern of each matrix by identifying a missing shape (from a multiple‐choice selection). The resulting raw score (number of correctly solved items) will be age-adjusted and converted to the standardized score.

*Verbal counting task*

The children will be asked to recite the count list, starting from one and counting as high as they can without an error, or until they reach 100. They will be stopped whenever they make an error. The score will be the highest number reached without an error (Geary et al., 2018).

*Give-a-number task*

In this task, children will be presented with a pile of 30 small balls with a diameter of approximately 2 cm. The experimenter will ask them to provide a specific number of balls, using this instruction “In this game, I will ask you to give me some of these balls. I will ask you a few times. Ready? Could you give me x balls? Please place them on my plate.”

We used titrated version (Following the adapted version of the task in Krajcsi, 2021 and Marchand et al., 2022), in which we split the task into three pseudorandomized blocks to avoid the predictive nature of an increasing series: i) 1, 3, 2, ii) 6, 4, 5, iii) 9, 7, 8. Within each block, each number would be asked twice, and if the accuracy for that number is 50% (e.g. the first response is wrong, but the second is correct), the number would be asked the third time. If a child fails two or three numbers in the block (for example, gives correct answers for number 4, but not for 6 and 5), the next block will not be asked.

For scoring, the average performance (i.e., the proportion of the correctly solved trials for each number) will be calculated for each child. A number is known if two criteria are met: First, the proportion of correct responses must be higher than 66%, which means at least two or three correct answers for a given number. The number knowledge is specified as the largest known number after which the first unknown number follows (e.g., number knowledge is 3 if the child solves numbers 1-3 but fails for number 4). If children’s measured number of knowledge is at least 5, they will be considered CP-knowers, otherwise, they will be considered subset-knowers.

**fNIRS experiment**

The experimental design consists of an auditory number word adaptation paradigm (Vogel et al., 2017). Children will hear deviant auditory number words and a deviant auditory non-number word that will be presented via speakers. Deviant auditory number words include the monosyllabic number words ‘four’ and ‘eight’, which differ by ½ and ¼ ratios from ‘two’ as the adaptation auditory number word (i.e., 2/4 and 2/8). The following ratio has been chosen following the adaptation signal recovery that suggests the increased ratio (when the ratio is calculated by dividing the larger number by the smaller number) is linked to better signal recovery in adaptation task in the brain areas associated with symbolic processing (Vogel et al., 2017; Ansari, 2008; Chiou et al., 2023; Nieder & Dehaene, 2009). The deviant auditory non-number word ‘rin’ (Figure 2A) will be considered as the control condition. Therefore, there will be three conditions: the deviant auditory number word ‘four’ (within the range of subset numbers), the deviant auditory number word ‘eight’ (within the range of cardinality numbers but beyond the range of subset numbers) and the deviant auditory non-number word ‘rin’.

The rationale for splitting the deviant auditory number words ‘four’ and ‘eight’ into two separate conditions is that CP-knowers are expected to process the conceptual difference between both deviant auditory number words and adaptation auditory number word ‘two’, while subset-knowers are expected to process the deviant auditory number word ‘four’ but not ‘eight’. The experiment will follow a block design. During each block, the adaptation auditory number word ‘two’ will be repeatedly presented and alternated by one of the conditions. There will be three runs in total, and each run will contain 4 blocks of one condition. The order of runs will be counterbalanced using a Latin square design across participants. The rationale for having each run include only one of the conditions is to avoid task switching which is a very cognitively demanding task for young children, and also to have a better signal-to-noise ratio (Rahimpour et al., 2023). There will be 4 blocks per condition, and in each block, there will be 5 trials of deviants (the type of deviant depends on the condition) and 5 trials of adaptation auditory number word ‘two’ (Figure 2B).

The presentation length of each trial will be around 400 ms but slightly varies as a function of pronunciation time. The trials are separated by 1200 ms silence. Therefore, the approximate length of one block will be approximately 16 seconds. The blocks are separated by jittered inter-block intervals of 15-17 seconds of silence (mean of 16 sec), during which adaptation number word ‘two’ will be repeatedly presented (Figure 2B), the total time per experiment, for both English and Russian versions, will be approximately 6 minutes plus a self-paced break between the runs. The children will be familiarised with the task procedure through verbal instruction and prompted to attend to the stimuli without active engagement to produce a response. In order to minimise common difficulties specific to this age group (reluctance to wear a cap, boredom from a monotonous task, increased movement), participants will watch a silent cartoon of their choice during the experiment. The trials for English-speaking children will be recorded by an English-native female, while the trials for Russian-speaking children will be recorded by a Russian-native female.

A diagram of a number of words

Description automatically generated

Figure 2: A) Trial types. The two deviant auditory number words differ by ½ and ¼ ratios from ‘two’. The deviant auditory non-number word ‘rin’ is considered as the control condition. B) The block design paradigm for stimuli with an example of trials in one of the blocks. There are 4 blocks in one run. Note that to improve the signal-to-noise ratio, there is one condition per run. Each stimulus takes approximately 400 ms, with an inter-stimulus interval of 1200 ms and a jittered inter-block interval between 15-17 seconds. Other conditions (deviant auditory number word ‘four’ and deviant auditory nonnumber word ‘rin’) follow the same procedure. C) Positions of the fNIRS optodes based on fNIRS Optodes’ Location Decider (fOLD) toolbox (Zimeo Morais et al., 2018). The red circles depict the emitters, and the blue circles depict the detectors.

**Procedure**

Each child will be measured individually in a light-attenuated room in the nursery, at their home, or in the Baby lab at the University of Surrey thus providing participants with a well-known, safe environment. Before the data collection, we will recruit parents/guardians through emails and flyers at the nursery. Parents/guardians can choose to communicate with us only via email or meet us at the nursery and discuss any questions regarding the data collection. Once the parent/guardian reads and commits their consent, they will be asked medical background check questions, demographic information, as well as socioeconomic status questionnaire and home-maths activity questionnaire. Parents/guardians can choose to submit filled-out questionnaires either in person at the nursery or via email. During data collection, we will first introduce the fNIRS device to the child and their guardian if they wish to be present during data acquisition. The child will be also asked for their verbal agreement. Thereafter, one experimenter starts testing the child with the non-verbal IQ test, the verbal counting, and the give-a-number task, while the second experimenter records the child’s responses.

The two experimenters will then prepare the fNIRS cap and place the optodes on the child’s head. During the fNIRS preparation and wrapping up the fNIRS experiment, the child will watch a cartoon on a laptop. The fNIRS experiment starts with a short practice. The practice consists of a continuous stream of the adaptation auditory number word ‘two’ repeatedly presented. Unlike regular blocks, the practice block will consist of two deviant auditory number words (‘four’ and ‘eight’), and an auditory non-number word (‘rin’ ), to give an idea of what can be heard.

Once the fNIRS is set up, children are instructed to listen to the auditory sequence and watch a silent cartoon of their choice. No other response is needed for the trials and no feedback will be given. The experiment will consist of three runs (each approximately 2 minutes) with a self-paced break in between. The entire experiment including both fNIRS recording and behavioural measurements will be approximately 30-45 minutes (depending on the child’s level of engagement). To increase motivation, children will be awarded stickers after completing each task and a toy at the end, regardless of their performance.

**fNIRS data acquisition and preprocessing**

A portable continuous-wave fNIRS device (Brite, Artinis Medical Systems BV, The Netherlands; Figure 2C) with 8 sources and 10 detectors with dual-wavelength (760 nm and 850 nm) near-infrared light will be used. It measures relative concentration changes in oxygenated (HbO), deoxygenated (HHb), and total (HbT) hemoglobin (Maki et al., 1995) following the modified Beer-Lambert law (Cope & Delpy, 1988), which represents the brain activation changes. The sampling rate will be 25 Hz with the standard detectors located at 30 mm. The optodes (i.e., sources and detectors) are arrayed in a lattice pattern over the bilateral frontoparietal cortex and embedded in a Neoprene head cap (Figure 2C). This configuration forms 18 measurement channels, defined as the area under each source-detector pair. The left- and right-hemisphere emitters will be respectively placed on P1, P3, P5, CP3, CP5 and P2, P4, P6, CP4, CP6 for measuring the left and the right parietal regions respectively, and on FC1, FC3, F3, F5 and FC2, FC4, F4, F6 for measuring the left and the right frontal regions respectively, following the international 10-10 system (Chatrian et al., 1985).

Like Xie et al. (2022), before starting the experiment and throughout the experiment, we will ensure that each ROI has at least two (or in the rare more challenging cases – one) viable channels by monitoring the status of data acquisition on-line via OxySoft (OxySoft, Artinis Medical Systems, Elst, The Netherlands). Preprocessing and activation analyses will be carried out using the NIRS Brain AnalyzIR toolbox (Santosa et al., 2018). First, channel quality will be assessed once again using the QT-NIRS toolbox using SCI threshold = 0.6, Q threshold = 0.4 and PSP threshold = 0.06 (Bulgarelli et al., 2023; Hernandez & Pollonini, 2020). The resulting quality of data (quantified by the number of channels per ROI included) will be reported for both groups and individually for each participant in a table form. At this stage, only the datasets with at least one viable channel per ROI will continue to the next part of the preprocessing stage. Depending on the quality of data, the number of data points for Hypothesis 1 and Hypothesis 2 may differ: e.g., if there are no viable channels in the right parietal area, but there is at least one viable channel in the left parietal area, the recording will be analysed in Hypothesis 1 (that requires only the signal from the left parietal area), but not in Hypothesis 2 (that requires bilateral parietal signal).

Data will be converted to changes in optical density and corrected for motion artefacts using the Temporal Derivative Distribution Repair (TDDR) method with enabled PCA (Fishburn et al., 2019). This method uses a robust regression approach to reduce the magnitude of large fluctuations (e.g., motion) in the signal while leaving small fluctuations (i.e., hemodynamics) intact. This makes the TDDR an appropriate artefact correction method in young children (e.g., Fishburn et al., 2019; Quiñones-Camacho et al., 2019). Next, we will apply the fourth-order Butterworth with a band-pass of 0.01-0.09 Hz. Signals will be then converted to HbO and HHb concentrations using the modified Beer-Lambert law (Delpy and Cope, 1997) with a partial path length correction of 0.1 (i.e., differential path length factor = 6 and partial volume factor = 60) for both wavelengths. As the last step, we will apply systemic artefact correction using a general linear model (GLM), after which we will extract residuals that will be used in further processing.

Task activation will be quantified by convolving the boxcar function for each of the 12 blocks (deviant auditory number word ‘four’, deviant auditory number word ‘eight’ and deviant auditory non-number word ‘rin’) with the canonical hemodynamic response function (HRF) and submitting it to the first-level GLM. Slow drift in the signal will be corrected by including a third-order Legendre polynomial regressor in the GLM matrix (e.g., Fishburn et al., 2019). Coefficients will be estimated using the autoregressive pre-whitening approach using iteratively reweighted least squares (AR-IRLS), which controls for Type-I error in the fNIRS statistical model (Barker et al., 2013; Santosa et al., 2018, Santosa et al., 2017). This approach accounts for the serially correlated noise of systemic physiological oscillations (e.g., Fishburn et al., 2019). The estimated beta coefficients for all conditions will be extracted. The highest coefficient of HbO or the lowest coefficient of HHb for each condition on each region of interest (ROIs; bilateral parietal regions) will be used to test whether CP-knowers exhibit higher bilateral parietal activation, defined by higher HbO or lower HHb, specifically in the left parietal region, relative to subset-knowers (Hypothesis 1). This difference is expected in the contrast of deviant auditory number word ‘eight’ minus control (Hypothesis 1a) and in the contrast of deviant auditory number word ‘four’ minus control (Hypothesis 1b). Note that neurophysiologically, increased HbO or decreased HHb represent increased brain activation (Sholkmann et al., 2014). Therefore, we must expect opposite directions in HbO and HHb changes, in a way that if one goes up, the other goes down (Sholkmann et al., 2014). Accordingly, we will check HbO and HHb changes in each channel. The channels will be excluded if both HbO and HHb significantly increase or decrease as they are most probably contaminated by uncorrected artefact.

To calculate the functional connectivity between the ROIs, the autoregressive whitened signal of the channels with the highest beta coefficient for each condition and each ROI will be used. The AR correlation using robust regression between these signals will be calculated for each condition (auditory number word and auditory non-number word). The robust correlation coefficients of each condition will be calculated and quantified using the Fisher z-transformation of their absolute values (e.g., Fishburn et al., 2018). The Fisher z-transformed values of HbO will be used to test whether CP-knowers will exhibit higher bilateral parietal functional connectivity, defined by functional co-activation between these two regions, relative to subset-knowers (Hypothesis 2). This difference is expected in the contrast of deviant auditory number word ‘eight’ minus control (Hypothesis 2a) and the contrast of deviant auditory number word ‘four’ minus control (Hypothesis 2b). Statistical analysis will be conducted using R (R Core Team, 2022). The preprocessing and analysis scripts will be available on OSF (<https://osf.io/ntprc/>)

**Analyses plan**

We will apply the Bayesian approach to test our hypotheses (Soltanlou et al., 2019). The statistical analysis (Bayesian t-tests) will be performed using R (R Core Team, 2022). Since this is the first study of its kind and we cannot infer priors from the literature at present, we will use uninformed Cauchy priors for all Bayesian analyses, and models will be compared to the null model. For the t-tests, the Cauchy priori scale will be set to a default of 0.707 (Quintana & Williams, 2018). For the interpretation, we will rely on calculating Bayes Factors (BF), the likelihood of the proposed alternative hypotheses (H1) compared to the null hypotheses (H0). The Bayes factor is typically denoted as BF01 when indicating the likelihood of the H0 over H1 and as BF10 when indicating the likelihood of the H1 over H0 (Schönbrodt & Wagenmakers, 2018). We adopted the scheme by Lee and Wagenmakers (2013) to describe the levels of evidence. When BF10 is between 1 and 3, it corresponds to anecdotal evidence in favour of H1 over H0; when BF10 is between 3 and 10, it corresponds to moderate evidence; lastly, when BF10 is higher than 10, it corresponds to strong evidence.

To test for Hypothesis 1, we will run two-tailed Bayesian independent *t*-tests comparing parietal engagement for experimental conditions between CP-knowers and subset-knowers. We will compare brain activation of CP-knowers and subset-knowers in the left parietal area for the contrast of the deviant auditory number word ‘eight’ – control (Hypothesis 1a) and for the contrast of the deviant auditory number word ‘four’ – control (Hypothesis 1b). We summarised the detailed prediction and rationale for each hypothesis in Table 1.

Table 1. Predicted results for Hypothesis 1

|  |  |  |  |
| --- | --- | --- | --- |
| **Hypothesis 1** | **Type of analysis** | **Expected result** | **Rationale** |
| Hypothesis 1a:  CP-knowers will have higher activations in the **left parietal** area than subset-knowers when comparing the contrast of deviant auditory number word ‘eight’ – control. | Independent two-tailed t-test between CP-knowers and subset-knowers | CP-knowers will have higher left parietal activation when preprocessing the word ‘eight’. | This prediction is based on the assumption that CP-knowers already understand the concept of the number word ‘eight’, unlike subset-knowers, thus the engagement of the left parietal region responsible for the numerical processing will be more evident. |
| Hypothesis 1b.  CP-knowers will have higher activations in the **left parietal** area than subset-knowers when comparing the contrast of deviant auditory number word ‘four’ – control. | Independent two-tailed t-test between CP-knowers and subset-knowers | CP-knowers will have higher  left parietal activation when  preprocessing the word  ‘four’. | This prediction is based on the assumption that while both CP- and subset-knowers understand the semantic meaning of the number word ‘four’ and will show left parietal activation in response, in CP-knowers, due to their advancement, the processing of the number word ‘four’ in the left parietal area will be more evident than in subset-knowers. |

Note that in the contrast comparison, the deviant auditory non-number word is referred to as control.

To test for Hypothesis 2, we will run two-tailed Bayesian independent *t*-tests comparing parietal engagement for experimental conditions between CP-knowers and subset-knowers. We will compare the connectivity between the left and right parietal areas both for the contrast of the deviant auditory number word “eight” – control (Hypothesis 2a), and for the contrast of the deviant auditory number word “four” – control (Hypothesis 2b). We summarised the detailed prediction and rationale for each sub-hypothesis in Table 2.

Table 2. Predicted results for Hypothesis 2

|  |  |  |  |
| --- | --- | --- | --- |
| **Hypothesis 2** | **Type of analysis** | **Expected result** | **Rationale** |
| Hypothesis 2a:  CP-knowers will have higher functional connectivity between **left parietal and right parietal** areas than subset-knowers when comparing the contrast of deviant auditory number word ‘eight’ – control. | Independent two-tailed t-test between CP-knowers and subset-knowers | CP-knowers will have higher bilateral parietal connectivity when preprocessing the word ‘eight’. | This prediction is based on the assumption that CP-knowers already built a link between the number word ‘eight’, and its intuitive representation, unlike subset-knowers that cannot process this number word yet. |
| Hypothesis 2b:  CP-knowers will have higher functional connectivity between **left parietal and right parietal** areas than subset-knowers when comparing the contrast of deviant auditory number word ‘four’ – control. | Independent two-tailed t-test between CP-knowers and subset-knowers | CP-knowers will have higher  bilateral parietal connectivity  when preprocessing the word  ‘four’. | This prediction is based on the assumption that while both CP- and subset-knowers understand the semantic meaning of the number word ‘four’ and will show parietal engagement in response, CP-knowers, due to their advancement, built a stronger link between the number word ‘four’, and its intuitive representation, unlike subset knowers that have just started processing this number word. |

Note that in the contrast comparison, the deviant auditory non-number word is referred to as control.

Please note, Hypotheses 1 and 2 are expected to be primarily driven by strong evidence for Hypotheses 1a and 2a (related to the number word 'eight'), whereas Hypotheses 1b and 2b (related to the number word 'four') will likely provide weaker evidence as compared to Hypotheses 1a and 2a. A smaller group difference is anticipated for the number word 'four' because while both CP- and subset-knowers understand the semantic meaning of the number word ‘four’ and are expected to show a strong parietal response, CP-knowers will show a stronger parietal response than subset-knowers due to their advancement. In contrast, a larger group difference is anticipated for the number word 'eight' because subset-knowers, unlike CP-knowers, do not yet understand the number word 'eight' and are not expected to show a strong parietal response.

**References**

Ansari, D. (2008). Effects of development and enculturation on number representation in the brain.

*Nature reviews neuroscience*, *9*(4), 278-291.

Ansari, D., Garcia, N., Lucas, E., Hamon, K., & Dhital, B. (2005). Neural correlates of symbolic number processing in children and adults. *Neuroreport*, *16*(16), 1769-1773.

Arsalidou, M., Pawliw-Levac, M., Sadeghi, M., & Pascual-Leone, J. (2018). Brain areas associated with numbers and calculations in children: Meta-analyses of fMRI studies. *Developmental cognitive neuroscience*, *30*, 239-250.

Arsalidou, M., & Taylor, M. J. (2011). Is 2+ 2= 4? Meta-analyses of brain areas needed for numbers and calculations. *Neuroimage*, *54*(3), 2382-2393.

Aslin, R. N., Shukla, M., & Emberson, L. L. (2015). Hemodynamic correlates of cognition in human infants. *Annual review of psychology, 66*, 349-379.

Barker, J. W., Aarabi, A., & Huppert, T. J. (2013). Autoregressive model based algorithm for correcting motion and serially correlated errors in fNIRS. *Biomedical optics express*, *4*(8), 1366-1379.

Barratt, W. (2006). The Barratt simplified measure of social status (BSMSS). Terre Haute, IN: Indiana State University

Barreto, C., Soltanlou, M. (2022). Functional near-infrared spectroscopy as a tool to assess brain activity in educational settings: An introduction for educational researchers. South African Journal of Childhood Education. 12(1), a1138. doi:10.4102/sajce.v12i1.1138

Ben-Shalom, T., Berger, A., & Henik, A. (2013). My brain knows numbers!-an ERP study of preschoolers’ numerical knowledge. *Frontiers in psychology*, *4*, 716.

Berger, A., Shmueli, M., Lisson, S., Ben-Shachar, M. S., Lindinger, N. M., Lewis, C. E., ... & Jacobson, S. W. (2019). Deficits in arithmetic error detection in infants with prenatal alcohol exposure: An ERP study. *Developmental cognitive neuroscience*, *40*, 100722.

Berger, A. (2011). Electrophysiological evidence for numerosity processing in infancy. *Developmental neuropsychology*, *36*(6), 668-681.

Bugden, S., Park, A. T., Mackey, A. P., & Brannon, E. M. (2021). The neural basis of number word processing in children and adults. *Developmental Cognitive Neuroscience*, *51*, 101011.

Bulgarelli, C., Pinti, P., Aburumman, N., & Jones, E. J. (2023). Combining wearable fNIRS and immersive virtual reality to study preschoolers’ social development: a proof-of-principle study on preschoolers’ social preference. Oxford Open Neuroscience, 2.

Cantlon, J. F., Brannon, E. M., Carter, E. J., & Pelphrey, K. A. (2006). Functional imaging of numerical processing in adults and 4-y-old children. *PLoS biology*, *4*(5), e125.

Cantlon, J. F., Davis, S. W., Libertus, M. E., Kahane, J., Brannon, E. M., & Pelphrey, K. A. (2011). Inter-parietal white matter development predicts numerical performance in young children. *Learning and Individual Differences*, *21*(6), 672-680.

Cantlon, J. F., Libertus, M. E., Pinel, P., Dehaene, S., Brannon, E. M., & Pelphrey, K. A. (2009). The neural development of an abstract concept of number. *Journal of cognitive neuroscience*, *21*(11), 2217-2229.

Chatrian, G. E., Lettich, E., & Nelson, P. L. (1985). Ten percent electrode system for topographic studies of spontaneous and evoked EEG activities. *American Journal of EEG technology*, *25*(2), 83-92.

Chiou, R., Margulies, D., Soltanlou, M., Jefferies, B., Cohen Kadosh, R. (2023). Semantic cognition vs. numerical cognition: A topographical perspective. Trends in Cognitive Sciences. 27 (11): 993- 995. doi:10.1016/j.tics.2023.08.004

Cipora, K., & Soltanlou, M. (2021). Direct and conceptual replication in numerical cognition. *Journal of Numerical Cognition*, *7*(3), 240-247.

Cope, M., & Delpy, D. T. (1988). System for long-term measurement of cerebral blood and tissue oxygenation on newborn infants by near infra-red transillumination. *Medical and Biological Engineering and Computing*, *26*, 289-294.

De Keyser, L., Bakker, M., Rathé, S., Wijns, N., Torbeyns, J., Verschaffel, L., & De Smedt, B. (2020). No Association Between the Home Math Environment and Numerical and Patterning Skills in a Large and Diverse Sample of 5- to 6-year-olds. Frontiers in Psychology, 11.

Delpy, D. T., & Cope, M. (1997). Quantification in tissue near–infrared spectroscopy. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, *352*(1354), 649-659.

Duncan, G. J., & Magnuson, K. (2012). Socioeconomic status and cognitive functioning: moving from correlation to causation. *Wiley Interdisciplinary Reviews: Cognitive Science*, *3*(3), 377-386.

Edwards, L. A., Wagner, J. B., Simon, C. E., & Hyde, D. C. (2016). Functional brain organization for number processing in pre‐verbal infants. *Developmental science*, *19*(5), 757-769.

Elliot, C. D., & Smith, P. (2011). British Ability Scales-(BAS-3). London, United Kingdom: GL Assessment.

Emerson, R. W., & Cantlon, J. F. (2012). Early math achievement and functional connectivity in the fronto-parietal network. *Developmental Cognitive Neuroscience*, *2*, S139-S151.Zhang et al., 2019

Fanelli, D. (2011). The black, the white and the grey areas: Towards an international and interdisciplinary definition of scientific misconduct. Promoting research integrity in a global environment, 79-90.

Fishburn, F. A., Ludlum, R. S., Vaidya, C. J., & Medvedev, A. V. (2019). Temporal derivative distribution repair (TDDR): a motion correction method for fNIRS. *Neuroimage*, *184*, 171-179.

Gallistel, C. R., & Gelman, R. (1990). The what and how of counting.

Geary, D. C., vanMarle, K., Chu, F. W., Rouder, J., Hoard, M. K., & Nugent, L. (2018). Early Conceptual Understanding of Cardinality Predicts Superior School-Entry Number-System Knowledge. *Psychological science, 29*(2), 191-205.

Haman, M., Lipowska, K., Soltanlou, M., Cipora, K., Domahs, F., Nuerk, H.-C. (2023). The plural counts: Inconsistent grammatical number hinders numerical development in preschoolers – a crosslinguistic study. Cognition. 235: 105383. doi:10.1016/j.cognition.2023.105383

Hernandez, S. M., & Pollonini, L. (2020, April). NIRSplot: a tool for quality assessment of fNIRS scans. In Optics and the Brain (pp. BM2C-5). Optica Publishing Group

Holloway, I. D., Battista, C., Vogel, S. E., & Ansari, D. (2013). Semantic and perceptual processing of number symbols-evidence from a cross-linguistic fMRI adaptation study. *Journal of cognitive neuroscience, 25*(3), 388-400.

Hubbard, E. M., Diester, I., Cantlon, J. F., Ansari, D., Van Opstal, F., & Troiani, V. (2008). The evolution of numerical cognition: From number neurons to linguistic quantifiers. *Journal of Neuroscience*, *28*(46), 11819-11824.

Hyde, D. C., Boas, D. A., Blair, C., & Carey, S. (2010). Near-infrared spectroscopy shows right parietal specialization for number in pre-verbal infants. *Neuroimage*, *53*(2), 647-652.

Hyde, D. C., & Spelke, E. S. (2011). Neural signatures of number processing in human infants: evidence for two core systems underlying numerical cognition. *Developmental science*, *14*(2), 360-371.

Hyde, D. C. (2021). The emergence of a brain network for numerical thinking. *Child Development Perspectives*, *15*(3), 168-175.

Izard, V., Dehaene-Lambertz, G., & Dehaene, S. (2008). Distinct cerebral pathways for object identity and number in human infants. *PLoS biology*, *6*(2), e11.

Kleemans, T., Peeters, M., Segers, E., & Verhoeven, L. (2012). Child and home predictors of early numeracy skills in kindergarten. Early Childhood Research Quarterly, 27(3), 471–477.

Kovacs, G., & Schweinberger, S. R. (2016). Repetition suppression - An integrative view. *Cortex, 80*, 1-4.

Krajcsi, A. (2021). Follow-up questions influence the measured number knowledge in the Give-a-number task. *Cognitive Development*, *57*, 100968.

Le Corre, M., & Carey, S. (2007). One, two, three, four, nothing more: An investigation of the conceptual sources of the verbal counting principles. *Cognition*, *105*(2), 395-438.

Lee, M. D., & Wagenmakers, E. J. (2013). *Bayesian cognitive modeling: A practical course*. Cambridge university press.

LeFevre, J.-A., Skwarchuk, S.-L., Smith-Chant, B. L., Fast, L., Kamawar, D., & Bisanz, J. (2009). Home numeracy experiences and children’s math performance in the early school years. Canadian Journal of Behavioural Science/Revue Canadienne Des Sciences Du Comportement, 41(2), 55–66.

Lewis, M., & Mayes, L. C. (2012). The role of environments in development: An introduction. *The Cambridge handbook of environment in human development*, 1-12.

Maki, A., Yamashita, Y., Ito, Y., Watanabe, E., Mayanagi, Y., & Koizumi, H. (1995). Spatial and temporal analysis of human motor activity using noninvasive NIR topography. *Medical physics*, *22*(12), 1997-2005.

Marchand, E., Lovelett, J. T., Kendro, K., & Barner, D. (2022). Assessing the knower-level framework: How reliable is the Give-a-Number task?. Cognition, 222, 104998.

MATLAB. (2023). V*ersion R2023b*. Natick, Massachusetts: The MathWorks Inc.

Merkley, R., & Ansari, D. (2016). Why numerical symbols count in the development of mathematical skills: Evidence from brain and behavior. Current Opinion in Behavioral Sciences, 10, 14-20.

Nieder, A., & Dehaene, S. (2009). Representation of number in the brain. *Annual review of neuroscience*, *32*, 185-208.

Pinhas, M., Donohue, S. E., Woldorff, M. G., & Brannon, E. M. (2014). Electrophysiological evidence for the involvement of the approximate number system in preschoolers' processing of spoken number words. *Journal of cognitive neuroscience, 26*(9), 1891-1904.

Pinti, P., Aichelburg, C., Gilbert, S., Hamilton, A., Hirsch, J., Burgess, P., & Tachtsidis, I. (2018). A review on the use of wearable functional near‐infrared spectroscopy in naturalistic environments. Japanese Psychological Research, 60(4), 347-373.

Quiñones-Camacho, L. E., Fishburn, F. A., Camacho, M. C., Wakschlag, L. S., & Perlman, S. B. (2019). Cognitive flexibility-related prefrontal activation in preschoolers: A biological approach to temperamental effortful control. *Developmental cognitive neuroscience*, *38*, 100651.

Quintana, D. S., & Williams, D. R. (2018). Bayesian alternatives for common null-hypothesis significance tests in psychiatry: a non-technical guide using JASP. *BMC psychiatry*, *18*, 1-8.

Rahimpour, A., Lanka, P., Pollonini, L., Proksch, S., Balasubramaniam, R., Bortfeld, H. (2023). Multiple Levels of Contextual Influence on Action-Based Timing Behavior and Cortical Activation. *Scientific Report.*

Rathé, S., Torbeyns, J., De Smedt, B., & Verschaffel, L. (2020). Are children’s spontaneous number focusing tendencies related to their home numeracy environment? ZDM, 52(4), 729–742.

Rivera, S. M., Reiss, A. L., Eckert, M. A., & Menon, V. (2005). Developmental changes in mental arithmetic: evidence for increased functional specialization in the left inferior parietal cortex. *Cerebral cortex*, *15*(11), 1779-1790.

Santosa, H., Aarabi, A., Perlman, S. B., & Huppert, T. J. (2017). Characterization and correction of the false-discovery rates in resting state connectivity using functional near-infrared spectroscopy. *Journal of Biomedical Optics*, *22*(5), 055002-055002.

Santosa, H., Zhai, X., Fishburn, F., & Huppert, T. (2018). The NIRS brain AnalyzIR toolbox. *Algorithms*, *11*(5), 73.

Sarnecka, B. W., Goldman, M. C., & Slusser, E. B. (2015). How counting leads to children’s first representations of exact, large numbers. *The Oxford handbook of numerical cognition*, 291-309.

Sarnecka, B. W., Kamenskaya, V. G., Yamana, Y., Ogura, T., & Yudovina, Y. B. (2007). From grammatical number to exact numbers: Early meanings of ‘one’,‘two’, and ‘three’in English, Russian, and Japanese. Cognitive psychology, 55(2), 136-168.

Scholkmann, F., Kleiser, S., Metz, A. J., Zimmermann, R., Pavia, J. M., Wolf, U., & Wolf, M. (2014). A review on continuous wave functional near-infrared spectroscopy and imaging instrumentation and methodology. *Neuroimage, 85*, 6-27.

Schönbrodt, F. D., & Wagenmakers, E. J. (2018). Bayes factor design analysis: Planning for compelling evidence. *Psychonomic bulletin & review*, *25*(1), 128-142.

Simon, G., Lubin, A., Houdé, O., & De Neys, W. (2015). Anterior cingulate cortex and intuitive bias detection during number conservation. *Cognitive neuroscience*, *6*(4), 158-168.

Sokolowski, H. M., Fias, W., Mousa, A., & Ansari, D. (2017). Common and distinct brain regions in both parietal and frontal cortex support symbolic and nonsymbolic number processing in humans: A functional neuroimaging meta-analysis. *Neuroimage*, *146*, 376-394.

Soltanlou, M., Artemenko, C., Dresler, T., Haeussinger, F. B., Fallgatter, A. J., Ehlis, A.-C., Nuerk, H.-C. (2017). Increased arithmetic complexity is associated with domain-general but not domainspecific magnitude processing in children: A simultaneous fNIRS-EEG study. Cognitive, Affective, & Behavioural Neuroscience. 17(4), 724-736. doi:10.3758/s13415-017-0508-x

Soltanlou, M., Artemenko, C., Ehlis, A.-C., Huber, S., Fallgatter, A. J., Dresler, T., Nuerk, H.-C. (2018). Reduction but not shift in brain activation in arithmetic learning in children: A simultaneous fNIRS-EEG study. Scientific Reports. 8(1), 1707. doi:10.1038/s41598-018-20007-x

Soltanlou, M., Coldea, A., Artemenko, C., Ehlis, A.-C., Fallgatter, A. J., Nuerk, H.-C., Dresler, T. (2019). No difference in the neural underpinnings of number and letter copying in children: Bayesian analysis of functional near-infrared spectroscopy data. Mind, Brain, and Education. 13(4), 313- 325. doi:10.1111/mbe.12225

Soltanlou, M., Dresler, T., Artemenko, C., Rosenbaum, D., Ehlis, A.-C., Nuerk, H.-C. (2022). Training causes activation increase in temporo-parietal and parietal regions in children with mathematical disabilities. Brain Structure and Function. 227, 1757-1771. doi:10.1007/s00429- 022-02470-5

Soltanlou, M., Sitnikova, M. A., Nuerk, H.-C., & Dresler, T. (2018). Applications of functional near-Infrared spectroscopy (fNIRS) in studying cognitive development: The case of mathematics and language. Frontiers in psychology, 9, 277.

Stefan, A. M., Gronau, Q. F., Schönbrodt, F. D., & Wagenmakers, E. J. (2019). A tutorial on Bayes Factor Design Analysis using an informed prior. *Behavior research methods*, *51*, 1042-1058.

Team, R. C. (2022). R Core Team R: a language and environment for statistical computing. Foundation for Statistical Computing.

Vogel, S. E., Goffin, C., & Ansari, D. (2015). Developmental specialization of the left parietal cortex for the semantic representation of Arabic numerals: an fMR-adaptation study. *Developmental Cognitive Neuroscience, 12*, 61-73.

Vogel, S. E., Goffin, C., Bohnenberger, J., Koschutnig, K., Reishofer, G., Grabner, R. H., & Ansari, D. (2017). The left intraparietal sulcus adapts to symbolic number in both the visual and auditory modalities: Evidence from fMRI. *Neuroimage, 153*, 16-27.

Whiteman, A. C., Santosa, H., Chen, D. F., Perlman, S., & Huppert, T. (2017). Investigation of the sensitivity of functional near-infrared spectroscopy brain imaging to anatomical variations in 5- to 11-year-old children. *Neurophotonics, 5*(1), 011009.

Wynn, K. (1990). Children's understanding of counting. *Cognition*, *36*(2), 155-193.

Xie, S., Gong, C., Lu, J., Li, H., Wu, D., Chi, X., & Chang, C. (2022). Enhancing Chinese preschoolers’ executive function via mindfulness training: An fNIRS study. Frontiers in Behavioral Neuroscience, 16, 961797.

Zimeo Morais, G. A., Balardin, J. B., & Sato, J. R. (2018). fNIRS Optodes’ Location Decider (fOLD): a toolbox for probe arrangement guided by brain regions-of-interest. Scientific reports, 8(1), 1-11

Zhang, H., Wee, C. Y., Poh, J. S., Wang, Q., Shek, L. P., Chong, Y. S., ... & Qiu, A. (2019). Fronto-parietal numerical networks in relation with early numeracy in young children. *Brain Structure and Function*, *224*, 263-275.

The Study Design Template

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Question** | **Hypothesis** | **Sampling plan** | **Analysis Plan** | **Rationale for deciding the sensitivity of the test for confirming or disconfirming the hypothesis**  **hypothesis** | **Interpretation given different outcomes** | **Theory that could be shown wrong by the outcomes** |
| Does CP knowledge relate to activation of the left parietal region? | **Hypothesis 1:** CP-  knowers will exhibit higher bilateral parietal activation, specifically in the left parietal region, relative to subset-knowers because they are more advanced in their conceptual knowledge about the meaning of number words.  subset- | We applied Bayes Factor Design Analysis (BFDA) using sequential Bayes factor with maximal n design, in which participants will be collected until either a sufficient BF is achieved, or a maximum number is reached (Schönbrodt & Wagenmakers, 2018). The benefit of using this design is that it ensures collecting sufficient sample while ensuring feasibility (Schönbrodt &  Wagenmakers, 2018). Calculation of sample size involved running Monte Carlo simulations and relied on the web-based BFDA application (Stefan et al., 2019; see http://shinyapps.org/apps/BFDA/) Simulation results indicate that with informed prior distribution of effect sizes (t(mu = 0.35, r = 0.102, df = 3; Stefan et al., 2019), the expected sample size to reach BF10 of at least 6 is 46 participants per group. As the testing with 20 participants per group is highly likely to cause misleading evidence (see again Stefan et al., 2019; Schönbrodt et al., 2017), we will start consecutive testing for the hypotheses once there are 25 participants per group and will stop once the testing reaches BF **>** 6 in favour of H1 or H0 or the maximum of 46 participants per group has been collected. Given that a total of 92 participants would be considered as a large sample for a neuroimaging study in preschoolers, if the BF for either of the hypotheses does not reach at least 6 (whether due to the effect’s extreme weakness or the common in the field deficit in neuroimaging data quality in preschoolers) even after recruiting maximum feasible sample size, we will consider this inconclusive result as an important message to the field.  (Schönbrodt & Wagenmakers, | Independent two-tailed t-test between CP-knowers and subset-knowers  – see Analysis plan section in paper. | Since this is the first fNIRS  study of numerical knowledge acquisition in preschool children, sample size calculation is necessarily inexact. A relevant fMRI study of a similar numerical adaptation paradigm in adults reported Cohen’s d effect size of 0.73 (Holloway, Battista, Vogel, & Ansari, 2013).  However, considering also the file drawer problem of unpublish ed studies, and publication bias towards significant findings (Cipora & Soltanlou, 2021;  Fanelli, 2011), as  well as the current lack of studies on the neuroimaging differences between subset-knowers and CP-knowers, we have decided to utilize a low to medium effect size of  Cohen's d = 0.35.  –  see Material  and Methods  section in the  paper.  is | BF10 3-10 of difference in the left parietal region in CP-knowers compared to subset-knowers will be taken as moderate evidence and the BF10 1-3 of difference in the left parietal region in CP-knowers compared to subset-knowers will be taken as anecdotal evidence.  –  see  Analysis  plan  section in  paper.  strong | Failure to find  evidence for  higher  activation in the  left parietal  region in CP-knowers  compared  to subset-knowers might  suggest that the  emergence  of symbolic  numerical  knowledge may  not necessarily  rely on the  parietal network  in young  children.  Explanations for  all results will  be presented  in the discussion.  discussion.might |
| Does CP knowledge relate to the functional connectivity of the left and right parietal regions within the parietal network of mathematical thinking? | Hypothesis 2: CP-knowers will exhibit higher functional connectivity between the left and right parietal regions relative to subset-knowers because they have built a link between number words and intuitive representations of quantity. | See above. | See above. | See above. | See above. | See above. |
|  |  |  |  |  |  |  |