Foundational number sense training gains are predicted by hippocampal–parietal circuits

Hyesang Chang, Lang Chen, Yuan Zhang, Ye Xie, Carlo de Los Angeles, Emma Adair, Gaston Zanitti, Demian Wassermann, Miriam Rosenberg-Lee, Vinod Menon

Foundational number sense training gains are predicted by hippocampal–parietal circuits

Abbreviated title: Learning related hippocampal–parietal circuits

Hyesang Chang[1], Lang Chen[1,2], Yuan Zhang[1], Ye Xie[1,3,4], Carlo de Los Angeles[1], Emma Adair[1], Gaston Zanitti[5], Demian Wassermann[5], Miriam Rosenberg-Lee[1,6], Vinod Menon[1,7,8*]*

[1]Department of Psychiatry & Behavioral Sciences, Stanford University School of Medicine, Stanford, CA 94305, US

[2]Department of Psychology, Santa Clara University, Santa Clara, CA 95053, US

[3]Department of Physics, Zhejiang University, Hangzhou 310027, China

[4]Department of Psychology, Sun Yat-Sen University, Guangzhou 510006, China

[5]Parietal, Inria Saclay Île-de-France, CEA Université Paris Sud 1 Rue Honoré d'Estienne d'Orves 91120, Palaiseau, France

[6]Department of Psychology, Rutgers University, Newark, NJ 07102, US

[7]Department of Neurology & Neurological Sciences,

[8]Stanford Neurosciences Institute,

Stanford University School of Medicine, Stanford, CA 94305, US

*Hyesang Chang, Ph.D.; Vinod Menon, Ph.D.

401 Quarry Rd., Stanford, CA 94305, US

Email: changh@stanford.edu; menon@stanford.edu

Number of pages: 65

Number of figures: 7

Number of tables: 10

Number of words for abstract: 246

Number of words for introduction: 650

Number of words for discussion: 1487

Conflict of Interest Statement: The authors declare no competing financial interests.

Acknowledgements

This research was supported by the United States National Institutes of Health to V.M. (HD059205, MH084164, HD094623) and M.R.-L. (MH101394), and by the Stanford Maternal & Child Health Research Institute Postdoctoral Support Award to H.C. We thank participating families, Sarit Ashkenazi, Laxman Dhulipala, Dawlat El-Said, Teresa Iuculano, Dietsje Jolles, Shelby Karraker, Samantha Mitsven, Sangeetha Santhanam, and Kaustubh Supekar for assistance with the study and Booil Jo and Jane P. Kim for advice on statistical analysis. We also thank Kristen Pilner Blair for assistance with the development of the Restaurant Game as part of number sense training program and Valentin Iovene for contributing to the development of reverse meta-analysis tool.

Abstract

The development of mathematical skills in early childhood relies on number sense, the foundational ability to discriminate between quantities. Number sense in early childhood is predictive of academic and professional success, and deficits in number sense are thought to underlie lifelong impairments in mathematical abilities. Despite its importance, the brain circuit mechanisms that support number sense learning remain poorly understood. Here, we designed a theoretically motivated training program to determine brain circuit mechanisms underlying foundational number sense learning in female and male elementary school-aged children (ages 7-10). Our four-week integrative number sense training program gradually strengthened the understanding of the relations between symbolic (Arabic numerals) and non-symbolic (sets of items) representations of quantity. We found that our number sense training program improved symbolic quantity discrimination ability in children across a wide a range of math abilities including those with learning difficulties. Crucially, the strength of pre-training functional connectivity between the hippocampus and intraparietal sulcus, brain regions implicated in associative learning and quantity discrimination, respectively, predicted individual differences in number sense learning across typically developing children and children with learning difficulties. Reverse meta-analysis of inter-regional co-activations across 14,371 fMRI studies and 89 cognitive functions confirmed a reliable role for hippocampal–intraparietal-sulcus circuits in learning. Our study identifies a canonical hippocampal–parietal circuit for learning which plays a foundational role in children’s cognitive skill acquisition. Findings provide important insights into neurobiological circuit markers of individual differences in children’s learning and delineate a robust target for effective cognitive interventions.

Significance Statement

Mathematical skill development relies on number sense, the ability to discriminate between quantities. Here, we develop a theoretically motivated training program and investigate brain circuits that predict number sense learning in children during a period important for acquisition of foundational cognitive skills. Our integrated number sense training program was effective in children across a wide a range of math abilities, including children with learning difficulties. We identify hippocampal–parietal circuits that predict individual differences in learning gains. Our study identifies a novel brain circuit predictive of the acquistion of foundational number sense skills and delineates a robust target for effective interventions and monitoring response to cognitive training.

Introduction

Number sense, the ability to discriminate quantities, is predictive of academic achievement and professional success (Jordan, Kaplan, Ramineni, & Locuniak, 2009; National Mathematics Advisory Panel, 2008). In particular, weaknesses in mapping symbolic numbers (e.g., the symbol “2”) to their magnitude representations (e.g., two objects) in early childhood are associated with difficulties in subsequent mathematical skill acquisition (De Smedt & Gilmore, 2011; Rousselle & Noël, 2007). The knowledge about the neural mechanisms that support number sense acquisition in children can provide important insights into neurobiological markers of individual differences in learning and inform more effective interventions. Here we develop a theoretically motivated training protocol to address a critical gap in our understanding of the acquisition of foundational skills and the brain circuits that predict learning in early elementary school children.

Recent studies have begun to uncover a crucial role for the hippocampal memory system in mathematical skill acquisition (Menon & Chang, 2021; Supekar, Chang, Mistry, Iuculano, & Menon, 2021). Most previous studies of mathematical cognition have focused on cortical regions, most notably the posterior parietal cortex; consequently, the role of medial temporal lobe regions in mathematical cognition and learning has received less attention because of this cortico-centric focus. Notably, in a study on arithmetic fact learning (e.g., 3 + 5 = 8), hippocampal functional connectivity with multiple cortical regions predicted individual differences in performance gains in children (Supekar et al., 2013). Furthermore, the strength of this association with learning was stronger than the connectivity of the intraparietal sulcus (IPS), a brain region consistently implicated in representation of numerical quantities across symbolic and non-symbolic formats (Butterworth & Walsh, 2011; Piazza & Eger, 2016). These findings provide support for theoretical models which posit that hippocampal circuitry is critical during early stages of math skill acquisition (Menon, 2016; Menon & Chang, 2021). Here we test the hypothesis that brain circuits linking the hippocampus, a brain region crucial for binding new information (Eichenbaum, 2004; Olsen, Moses, Riggs, & Ryan, 2012; Zeithamova & Bowman, 2020), and parietal cortical areas important for representation of numerial quantity supports foundational number sense learning.

We developed a number sense training protocol which emphasized strengthening children's understanding of the relations between symbolic and non-symbolic representations of quantity (**Figures 1A-B**). To probe learning during an important period for building foundational cognitive skills, we recruited 96, 7-10-year-old, children with a broad range of math abilities. Functional, diffusion, and structural MRI scans acquired before training and training-related changes in performance on symbolic quantity discrimination were used to determine integrity of brain circuits associated with learning in response to number sense training (**Figures 1C-E**). We had four goals. Our first goal was to determine whether number sense training is effective in children, including typically developing (TD) children and those with mathematical learning difficulties (MLD). Our second goal was to identify functional brain circuits that predict children's gains in number sense following training. We used task-independent, resting-state functional MRI to measure intrinsic functional connectivity, which is thought to reflect integrity of functional circuitry (Greicius, Krasnow, Reiss, & Menon, 2003) and is considered to be a relatively stable measure compared to task-dependent fMRI. Our third goal was to determine, as an adjunct to intrinsic functional connectivity measures, whether number sense training gains are predicted by white matter pathways linking the hippocampus with the parietal cortex. We used advanced High Angular Resolution Diffusion Imaging (HARDI) protocol that examines complex fiber tracts (Tuch et al., 2002) to enable high-quality reconstructions of white matter pathways. The final goal of our study was to expand on findings from our training study to investigate the broader role of hippocampal–parietal circuits in learning across a large set of fMRI studies using reverse meta-analysis (see *Methods*). We examined co-activations of hippocampus and parietal cortex across 14,371 published fMRI studies in relation to 89 cognitive functions to determine whether hippocampal–parietal functional circuits identified in the present study constitute a canonical circuit for learning.

Materials and Methods

Participants

A total of 96 children in 2nd and 3rd grades (age: *M* = 8.19, *SD* = 0.63, 54 females) recruited with flyers sent to schools and posted at libraries and community centers in the San Francisco Bay Area participated in the study. All participants were right-handed and did not report any current neurological or psychiatric illness. Among them, 66 children (age: *M* = 8.15, *SD* = 0.65, 35 females) participated in the training program and 30 children (age: *M* = 8.27, *SD* = 0.61, 19 females) served as no-contact controls. MLD in children was identified using *normed-based cutoff criteria* applied to math fluency, similar to previously published studies (Iuculano et al., 2015; Jolles et al., 2016; Rosenberg-Lee et al., 2015). Children who scored at or below 90 (25th percentile or below) on the Math Fluency subtest of the Woodcock Johnson Third Edition (WJ-III) (Woodcock, McGrew, & Mather, 2001) were included in the MLD group (*M* = 85.61, *SD* = 3.97), and children who scored above 90 were included in the TD group (*M* = 103.65, *SD* = 9.09). No participant was excluded due to MLD or TD status.

All study protocols were approved by the Stanford University School of Medicine Institutional Review broad and informed consent was obtained from the parents of the children. Children received $50 for completing each MRI scanning session, $50 for completing neuropsychological assessment battery, and $50 for participating in the training program. Twenty-seven children were excluded from resting-state fMRI analysis due to no structural MRI data acquired (*n* = 3), poor structural MRI image quality (*n* = 6), missing behavioral data from fMRI task (*n* = 2), or inadequate whole brain coverage or excessive head movement (*n* = 16; see fMRI preprocessing in *Methods*) in the scanner. For diffusion MRI analysis, 13 children were additionally excluded due to incomplete diffusion MRI data acquisition (n = 1), poor diffusion MRI data quality (n = 10), or identification as extreme outliers in structural connectivity measures (n = 2). The final resting-state fMRI analysis sample included 52 children (18 children with MLD, 34 TD children) and 17 children in the training and control group, respectively; the diffusion MRI analysis sample included 43 children (15 children with MLD, 28 TD children) and 13 children in the training and control group, respectively.

Experimental design and statistical analyses

The current study examined the brain circuit mechanisms that predict the acquisition of foundational cognitive skills, following integrative number sense training. The overall study design is summarized in **Figure 1A**. Children completed MRI scanning session and cognitive assessments before and after training (in the training group) or no contact (in the control group). The no-contact control group participated in all aspects of the study except for training to control for normal "business-as-usual" schooling (Fuchs et al., 2009) and determine the specificity of brain circuits that predict gains in intervention.

Children in the training group completed a 4-week number sense training program (3 days/week, for approx. 60 minutes/day), which focused on strengthening of children's understanding of the relations between symbolic and non-symbolic representations of quantity ranging from 1 to 9. The first week of training began with a review of counting principles, followed by practice in enumeration and comparisons between non-symbolic quantities (sets of items or dot arrays) in the second week, comparisons between non-symbolic and symbolic (Arabic numbers) quantities in the third week, and finally, comparisons between symbolic quantities in the fourth week. Training with non-symbolic numbers included in weeks 2 and 3 was used to scaffold children's learning of symbolic numbers through mapping non-symbolic quantities to verbal (number words) or visual (Arabic numerals) symbolic numbers. This was followed by training with symbolic numbers, without the use of non-symbolic numbers, in the final week. Response to training was examined using symbolic quantity discrimination task acquired before and after training.

For statistical analyses of behavioral data, two-sample *t*-tests, Chi squared tests, and multivariate analysis of variance (MANOVA) were performed for comparisons between groups of age, gender, or neuropsychological assessments. A repeated measures ANOVA with time as a within-subject factor and group as a between-subject factor was conducted to assess the effects of training on symbolic quantity discrimination task performance. Follow-up paired *t*-tests examined changes in task performance (learning gains) in each group. In addition, two sample *t*-tests assessed differences in learning gains and pre- and post-training differences in task performance between groups. Spearman correlations were used for analysis on relations between behavioral measures and brain-behavior relations to minimize influence of potential outliers. In addition to frequentist statistics (e.g., *p*-values), Bayes factor (BF) was used to assess presence or absence of evidence for H_1 or H_o (Keysers, Gazzola, & Wagenmakers, 2020). BF values greater than 3 provide evidence for H_1. BF values between .33 and 3 provide absence of evidence. BF values below .33 provide evidence of absence (evidence for H_o).

Details on statistical analyses of brain imaging data are described in *Intrinsic functional connectivity analysis*, *Structural connectivity analysis*, and *Cross-validation analysis* subsections. In addition, *Reverse meta-analysis* subsection describes the procedures for identifying cognitive functions associated with inter-regional coactivations reported in fMRI studies.

Training sessions

Across four weeks, children in the training group completed a variety of activities with a tutor (**Figure 1B**; **Table 1**). Generally, in each training session, children received a lesson on counting or comparisons, played computerized and interactive games, and completed review worksheets (counting or comparisons). Quantities from 1 to 9 were used in all activities. Upon successful completion of each activity, children added to a sticker sheet.

Lessons. In tutoring sessions in week 1, the tutor provided a lesson on counting by reviewing counting principles with examples of correct and incorrect counting of erasers; after the lesson, children viewed a video of a sock puppet counting and were asked to determine whether or not the sock puppet counted correctly. In tutoring sessions from weeks 2 to 4, children were asked to count out loud from 1 to 9 and was reminded that each number they counted up was bigger than the number before it. Then, in weeks 2 and 3, children completed *Math Circles*, where they enumerated the number (of erasers) in each of two *Math Circles* on the table and determined which number is bigger than the other. In week 2 (lesson on non-symbolic quantities), the tutor put two sets of erasers of different quantities in the two *Math Circles* (one quantity in each Math Circle). In week 3 (lesson on non-symbolic and symbolic quantities), the tutor put a card showing a number (Arabic numeral) and a set of erasers of different quantities from the number on the card in the two *Math Circles*. In week 4 (lesson on symbolic quantities), the tutor administered a number-ordering version of *Beat Your Score* (Chang, Rosenberg-Lee, Qin, & Menon, 2019), where the child was asked to order 4 decks of cards with quantities in non-symbolic (array of dots), mixed (symbolic and non-symbolic), or symbolic (Arabic numeral) format more quickly each time they ordered the cards. For each deck of cards, children completed the ordering of cards 3 times (after the tutor shuffled the cards), with the goal of beating their previous time taken to order the cards on the 3rd time for at least 3 decks of cards.

Games. In week 1, children played *Restaurant Game* (Blair, 2013a, 2013b) (computerized game), where they enumerated the number of dishes to cook for presented number of animals. From weeks 2 to 4, children played an adapted version of *Number Race* (Wilson, Revkin, Cohen, Cohen, & Dehaene, 2006) (computerized game), which was structured to align with the progression of training activities (comparison between non-symbolic quantities [week 2], symbolic and non-symbolic quantities [week 3], and symbolic quantities [week 4]) and did not include arithmetic training. Children also played *Math War* (Iuculano et al., 2015) (interactive game with the tutor) where they determined the larger quantity between two sets of dot arrays (week 2), between one set of dot arrays and a number (week 3), and between two numbers (week 4). In each *Math War* game, the child and the tutor each had a deck of cards, from which they drew one card at a time. The child first wrote down the number on their card and the number on the tutor's card and then marked the larger number. The game continued until the child and tutor drew all their cards. Finally, children played *Comparing Speed* with the tutor, where the child and the tutor identified cards with quantities one value above or below the quantities on the cards on the table. The cards had quantities in non-symbolic (week 2), symbolic and non-symbolic (week 3), or symbolic (week 4) format. In each *Comparing Speed* game, the tutor first put 4 cards on the table with quantities of 4, 5, 6, and 7. Then, the child and tutor each took 5 cards from two decks of cards (one for the child, another for the tutor). The child and the tutor placed their cards on top of any of 4 cards on the table when their cards had quantities one value above or below the quantities on the cards on the table. The child and the tutor drew more cards from their decks, keeping up to 5 cards in their hands. The game continued until the child finished placing their deck of cards and "won" the game.

Review worksheets. The review worksheets consisted of counting the number of animals (week 1) or identifying the larger quantity between two non-symbolic (week 2), symbolic and non-symbolic (week 3), or symbolic (week 4) quantities. Children circled, matched, or wrote the number of animals (week 1) or circled the larger quantity (weeks 2-4) on the worksheet. Accuracy was emphasized in the worksheets from weeks 1 to 3 (worksheet included 42 trials in week 1, 24 trials in week 2, and 48 trials in week 3). In week 4, children were given one minute to complete 24 symbolic number comparison trials the worksheet.

Symbolic quantity discrimination task

Before and after training (or no contact), children performed one run of symbolic quantity discrimination task in the MRI scanner, in which they had to determine which of two symbolic numbers presented on the screen was larger (**Figure 2A**). A total of 64 comparison trials, including quantities 1 through 9 (excluding 5) were presented in each run. Participants were instructed to press a left button if the left side had a larger quantity and the right button otherwise. Half of the trials had a near distance (1 unit) between the two quantities (e.g., 7:6), while the remaining trials had a far distance (5 units) between the two quantities (e.g., 3:8). Numerical magnitude was matched between the two distance conditions with an equal distribution of “big” (sum of pair of quantities greater than 10) and “little” (sum of pair of quantities less than 10) conditions. Our main outcome measure, number sense learning, was assessed by gains in performance efficiency in the symbolic quantity discrimination task. This was measured by the difference in symbolic quantity discrimination task efficiency (Townsend & Ashby, 1978), obtained by accuracy divided by reaction time, from pre- to post-training, with higher scores representing greater efficiency gains.

MRI data acquisition

Images were acquired on a 3T GE Signa scanner (General Electric, Milwaukee, WI) using a custom-built head coil at the Stanford University Lucas Center. Head movement was minimized during the scan by cushions placed around the participant’s head. A total of 31 axial slices (4.0 mm thickness, 0.5 mm skip) parallel to the anterior commissure (AC)-posterior commissure (PC) line and covering the whole brain were imaged using a T2* weighted gradient echo spiral in-out pulse sequence (Glover & Lai, 1998) with the following parameters: TR = 2 sec, TE = 30 msec, flip angle = 80°, 1 interleave. The field of view was 22 cm, and the matrix size was 64 x 64, providing an in-plane spatial resolution of 3.4375 mm. The total length of the run was 6 minutes and 10 seconds. To reduce blurring and signal loss from field inhomogeneity, an automated high-order shimming method based on spiral acquisitions was used before acquiring functional MRI scans (Kim, Adalsteinsson, Glover, & Spielman, 2002). High-resolution T1-weighted 3D MRI sequences were acquired to facilitate anatomical co-registration of fMRI maps, with the following parameters: I = 400ms, TR = 5.9ms; TE = minimum; flip angle = 11°; field of view = 240mm; matrix size = 256 × 192; 170 axial slices (1.0-mm thickness).

A state-of-the-art diffusion-weighted single-shot spin-echo, echo planar imaging HARDI pulse sequence was used for more precise examination of white-matter fibers, including crossing fibers (Tuch et al., 2002), with the following parameters: TR = 5.3s; TE = minimum; flip angle = 90°; field of view = 260mm; matrix size = 128 × 128; 50 axial slices (2.9-mm thickness, no spacing). The high b value (2500s/mm^2) was obtained by applying gradients along 150 different diffusion directions.

fMRI preprocessing

Resting-state functional MRI data were analyzed using SPM12 (http://www.fil.ion.ucl.ac.uk/spm/). The first 5 volumes were not analyzed to allow for T1 equilibration. A linear shim correction was applied separately for each slice during reconstruction (Glover & Lai, 1998). Images were realigned to the first scan to correct for motion and slice acquisition timing, co-registered to each individual's structural T1 images, spatially transformed to standard stereotaxic space (based on the Montreal Neurologic Institute coordinate system), resampled every 2 mm using sinc interpolation, and smoothed with a 6mm full-width half maximum Gaussian kernel to decrease spatial noise prior to statistical analysis. A bandpass filter (.008-.1 Hz) was applied to the smoothed data to remove high frequency artifacts. Translational movement in millimeters (x,y,z), and rotational motion in degrees (pitch, roll, yaw) were calculated based on the SPM12 parameters for motion correction of the functional images of each subject. We excluded participants with movement larger than 10mm in any of the x,y,z directions. Mean scan-to-scan displacement of movement did not exceed 0.5mm for all participants.

Intrinsic functional connectivity analysis

Intrinsic functional connectivity analysis was conducted using resting-state fMRI to investigate specific functional circuits that relate to change in symbolic quantity discrimination task performance with training. Regions of interest (ROIs) for functional connectivity were selected from Brainnetome (Fan et al., 2016) parcellations of the left and right hippocampus (rostral and caudal subdivisions combined). Seed-to-whole-brain functional connectivity for each ROI was estimated by extracting eigenvalues of time series of all voxels within the ROI, regressing out the global mean signal, white matter signal, cerebrospinal fluid signal, and six motion parameters. A bandpass filter (.008-.1 Hz) was applied to reduce high frequency noise. Functional connectivity maps were then submitted to a second-level analysis to examine whether the connectivity of these regions at the voxel-by-voxel level is predictive of learning (post-training – pre-training efficiency in symbolic quantity discrimination task).

Pre-training performance on symbolic quantity discrimination was regressed out to control for regression to the mean, a known phenomenon in intervention studies (Barnett, van der Pols, & Dobson, 2005), and to obtain more precise estimates of intervention effects (Pocock, Assmann, Enos, & Kasten, 2002; Thompson, Lingsma, Whiteley, Murray, & Steyerberg, 2015). This analysis approach allowed us to minimize the potential influence of pre-training performance on learning. In the current study, the association between change in performance and pre-training performance on symbolic quantity discrimination was significant in the training ($\rho = -.50$, $p < .001$) but not in the control ($\rho = -.18$, $p = .49$) group, which indicates that the degree to which pre-training performance influences changes in performance varied between groups.

Significant clusters were identified using a height threshold of $p < .005$ with multiple comparison correction at $p < .05$ after grey matter masking. This statistical threshold was chosen to balance between Type I and Type II errors in the current study, considering that larger sample sizes are typically needed to detect effects with a more stringent threshold (Carter, Lesh, & Barch, 2016). The cluster threshold was determined based on Monte Carlo simulations (Forman et al., 1995; Nichols & Hayasaka, 2003; Ward, 2000) implemented in custom Matlab scripts, similar to previous studies (Cho et al., 2012; Cho, Ryali, Geary, & Menon, 2011; Iuculano et al., 2015; Iuculano et al., 2014; Qin et al., 2014; Rosenberg-Lee, Barth, & Menon, 2011; Rosenberg-Lee et al., 2018). Ten thousand iterations of random 3D images, with the same resolution and dimensions as the fMRI data, were generated. The resulting images were masked for grey matter and then smoothed with the same 6mm FWHM Gaussian kernel used to smooth the fMRI data. The maximum cluster size was then computed for each iteration and the probability distribution was estimated across the 10,000 iterations. Based on this procedure, 67 voxels corresponding to $p < .05$ was used for the cluster threshold.

Follow-up ROI-based analyses were performed to visualize the results, ensure that the results were not driven by outliers, and confirm differences in correlation between brain and behavioral measures across groups. ROIs were defined as 6-mm spheres centered at peak of the left intraparietal sulcus (IPS) identified from hippocampal connectivity patterns in the training group and its contralateral region for the right IPS. Similar to the whole brain analysis (see above), pre-training symbolic quantity discrimination task efficiency was regressed out from symbolic quantity discrimination task efficiency gains and connectivity estimates in each group.

Structural connectivity analysis

Diffusion images were preprocessed to correct artifact issues from movement and eddy currents using FSL 5.0.11 (Andersson & Sotiropoulos, 2016). Then, probabilistic tractography to estimate structural connectivity between the left and right hippocampus and the left and right IPS ROIs was performed in native volume space using FSL's probtrackX (http://fsl.fmrib.ox.ac.uk/fsl/fslwiki/ (2007)). The hippocampus ROIs were from Brainnetome (Fan et al., 2016) parcellations (rostral and caudal subdivisions combined) and the IPS ROIs were from Brainnetome (Fan et al., 2016) parcellations that overlapped with target ROIs identified from intrinsic functional connectivity analysis (superior parietal lobule [lateral area 5] and inferior parietal lobule [rostrodorsal area 40] subdivisions combined; left IPS) and the contralateral regions (right IPS). These ROIs were warped to each subject's diffusion space, which was achieved by registering the B0 image of each subject's diffusion MRI image to MNI space using ANTs (Avants, Schoenemann, & Gee, 2006). ROIs were then dilated 3mm into the white matter to avoid biases generated by superficial white matter tracts (Thomas et al., 2014). Structural connectivity between two ROIs, A and B, was computed by the probability that diffusion images provide evidence that a white matter connection exist between these ROIs. This was calculated by the ratio between the number of tracts with origin in A or B reaching the other region and the total number of tracts seeded on A or B. Tracts were only considered if they stayed within the white matter and had a minimum length of 5mm between the ROIs. Finally, these measures were corrected for distance bias using the approach proposed by Donahue et al. (Donahue et al., 2016). To estimate structural connectivity between the ROIs, 5,000 tracts were seeded at each ROI for each individual, based on a preliminary analysis that determined the number of seeds needed to stabilize the connectivity measure. We computed structural connectivity strength between each pair of ROIs per subject. ROI-based analyses were then performed to assess the relation between hippocampal–parietal structural connectivity and learning, using similar procedures described in intrinsic functional connectivity analysis.

Cross-validation analysis

A machine-learning approach with balanced fourfold cross-validation combined with linear regression (https://github.com/poldrack/regressioncv) (Cohen et al., 2010; Supekar et al., 2013) was conducted to investigate the robustness of brain-based predictors of individual differences in number sense learning gains. Learning gain as a dependent variable and the brain-based predictor (connectivity) as an independent variable were treated as inputs to a linear regression algorithm. *r*(predicted, observed), a measure of how well the independent variable predicts the dependent variable, was first estimated using a balanced fourfold cross-validation procedure. Participants were assigned to one of four folds. A linear regression model was built using three folds leaving out the fourth, and this model was then used to predict the data in the left-out fold. This procedure was repeated four times to compute a final *r*(predicted, observed) representing the correlation between the data predicted by the regression model and the observed data. Finally, the statistical significance of the model was assessed using a non-parametric testing approach. The empirical null distribution of *r*(predicted, observed) was estimated by generating 1000 surrogate datasets under the null hypothesis that there was no association between changes in numerical skills and brain-based predictor.

Reverse meta-analysis

To examine the role of the hippocampal–parietal functional circuits, we conducted a novel reverse meta-analysis of inter-regional coactivation of the hippocampus and the parietal cortex, using regions identified in the present study, reported across 14,371 published fMRI studies up to July 2018 from NeuroSynth (Yarkoni, Poldrack, Nichols, Van Essen, & Wager, 2011) database and 89 cognitive atlas terms (CogAt) (Poldrack et al., 2011) (**Figure 7A**). The hippocampal ROIs were from Brainnetome (Fan et al., 2016) parcellations (rostral and caudal subdivisions combined) and the parietal ROIs were from Brainnetome (Fan et al., 2016) parcellations that overlapped with target ROIs identified from intrinsic functional connectivity analysis (superior parietal lobule [lateral area 5] and inferior parietal lobule [rostrodorsal area 40] subdivisions combined; left IPS) and the contralateral regions (right IPS).

Our reverse meta-analysis estimated the probability that a term related to a cognitive function was mentioned in an fMRI study, provided that activations in both the hippocampus and parietal cortex were also reported. For instance, we estimated the probability that for any given study:

P (term ‘learning’ is mentioned | activations are reported in the left hippocampus and in the left parietal cortex)

This probability was estimated across different domains and contexts in the neuroimaging literature. We performed this analysis on ipsilateral and contralateral hippocampal–parietal circuits on both hemispheres (i.e., left hippocampus – left IPS, right hippocampus – left IPS, left hippocampus – right IPS, right hippocampus – right IPS). To estimate these probabilities, we programmed this hypothesis in the probabilistic logic language NeuroLang (neurolang.github.io) (Iovene & Wassermann, 2020) using the full NeuroSynth (Yarkoni et al., 2011) open access database v0.7.

To estimate reverse meta-analysis probabilities, we followed the following steps. First, we encoded the probability of a term being present in a study by thresholding the term frequency – inverse document frequency (TF – IDF) value of the term being present at 10^{-3}, in agreement with NeuroSynth’s implementation (Yarkoni et al., 2011). Second, we considered the probability of a region being reported in a given study as directly proportional to the number of activations within the regions being present in the study, for which we resampled the activation foci to 4mm^3 voxels in MNI152 space. Third, terms were filtered using the CogAt (Poldrack et al., 2011) ontology to ensure that only those relating to cognitive processes (89 terms listed in **Figure 7B**) were taken into account. To assess the stability of our estimations, we computed the confidence interval of our reverse meta-analysis probability estimations, we split the 14,371 studies in 20 equal folds, maximizing the measurements for estimation. Finally, top 5% probable terms were considered to be sufficient evidence for associations with analyzed circuits.

Our analysis resulted in the selection of 25 out of 356 associations (4 circuits, 89 terms), which represented above 95 percentile of probable term mentions for studies where hippocampal–parietal circuits were reported. For these top 5% terms, the maximum probability was estimated, across all splits, at 0.34±0.011 for *memory* being mentioned in studies where left hippocampus and right IPS activations are simultaneously reported, and the minimum was estimated at 0.10±0.005 for mentioning *recognition* in studies where left hippocampus and left IPS activations are simultaneously reported. Hence, the standard deviation for all top 5% probabilities, is an order of magnitude smaller than the estimated probability, pointing out to a high confidence in our estimation.

Results

Comparison of neuropsychological measures between groups

A total of 96 children in 2nd and 3rd grades (age: M = 8.19, SD = 0.63, 54 females) participated in number sense training or served as no-contact controls (**Figure 1**). Sixty-nine of these children had high-quality behavioral and fMRI data (see *Methods* for details). We used two-sample *t*-tests, Chi squared tests, and multivariate analysis of variance (MANOVA) to compare age, gender, or neuropsychological assessments between the training and control groups.

Children in the training and control groups did not significantly differ in age ($t(67) = .66$, $p = .51$, Cohen's $d = .18$, BF = .33) and gender ($\chi^2_1 = .41$, $p = .52$, $\varphi = .08$, BF = .38; **Table 2**). A MANOVA between training and control groups on multiple neuropsychological assessments, including Wechsler Abbreviated Scale of Intelligence (WASI) (Wechsler, 1999) (Full-Scale, Verbal, and Performance IQ) and WJ-III (Math Fluency, Calculation, Applied Problems, Letter-Word Identification, and Word Attack) subtests ($F(8,60) = .56$, $p = .807$) showed no significant difference between groups. Two-sample *t*-tests confirmed that children in the training and control groups did not significantly differ in WASI ($|t|$s < .55, ps > .58, |Cohen's d| < .16, BFs < .33) and WJ-III subtests (ts < 1.31, ps > .19, Cohen's d < .37, BFs < .57; **Table 2**).

TD children and children with MLD in the training group were well-matched on age ($t(50) = 1.47$, $p = .15$, Cohen's $d = .43$, BF = .69) and gender ($\chi^2_1 = .01$, $p = .93$, $\varphi = .01$, BF = .34; **Table 3**). A MANOVA revealed a significant difference between TD and MLD groups on combined neuropsychological assessments ($F(8,43) = 8.09$, $p < .001$). TD children and children with MLD in the training group were well-matched on IQ measures ($|t|$s < 1.22, ps > .23, |Cohen's d| < .36, BF < .53; **Table 3**). As expected, children with MLD performed significantly worse than TD children on all of WJ-III math subtests ($|t|$s > 3.17, ps < .003, |Cohen's d| > .92, BFs > 14.57). Children with MLD also performed poorly on WJ-III reading subtests, compared to TD children ($|t|$s > 2.01, ps < .05, |Cohen's d| > .58), though there was insufficient evidence for group difference (.33 < BFs < 3).

In summary, children included in training and control groups were well-matched in terms of age, gender, and IQ, as well as other standardized measures of math and reading abilities. TD children and children with MLD in the training group were matched in age, gender, and IQ.

Compared to TD children, children with MLD performed poorly on both math and reading assessments, consistent with observations that comorbidity is one of the characteristics of MLD (Kaufmann & von Aster, 2012; Landerl, Göbel, & Moll, 2013). Nonetheless, strong evidence (BFs > 10) for group differences in math ability and insufficient evidence (.33 < BFs < 3) for group differences in reading ability indicate specific impairments in math skills in children with MLD. As shown in **Tables 4-5**, these results are similar for sample included in diffusion MRI data analysis (a subset of resting-state fMRI data analysis sample). In subsequent behavioral data analyses, we use the sample from resting-state fMRI data analysis.

Changes in performance on symbolic quantity discrimination in response to four weeks of number sense training

To assess children's behavioral performance on the symbolic quantity discrimination task (**Figure 2A**), we measured efficiency (Townsend & Ashby, 1978), derived from dividing accuracy by reaction time, to control for variations in speed-accuracy tradeoff and to reduce the number of statistical tests required. Higher efficiency scores indicated better performance. A repeated measures ANOVA on efficiency with time (pre/post) as a within-subject factor and group (training/control) as a between-subject factor was conducted to assess the effects of training on symbolic quantity discrimination task performance. Follow-up paired *t*-tests examined changes in task performance (learning gains) in each group and two sample *t*-tests assessed differences in learning gains and pre- and post-training differences in task performance between groups.

A repeated measures ANOVA on symbolic quantity discrimination task efficiency revealed a main effect of time ($F(1,67) = 29.37$, $p < .001$, $\eta_p^2 = .30$) and an interaction between time and group ($F(1,67) = 8.31$, $p = .005$, $\eta_p^2 = .11$). There was no significant main effect of group ($F(1,67)$

= .21, p = .65, η_p^2 < .01). Training significantly improved symbolic quantity discrimination task efficiency in the training group (t(51) = 6.28, p < .001, Cohen's d = .87, BF > 100), but not in the control group (t(16) = .17, p = .86, Cohen's d = .04, BF = .25, **Figure 2B**). Both TD children (t(33) = 5.47, p <.001, Cohen's d = .94, BF > 100) and children with MLD (t(17) = 3.16, p = .006, Cohen's d = .74, BF = 8.43) improved with large individual differences in both groups (coefficient of variation: TD children: 1.06; children with MLD: 1.34). In addition, learning gains – changes (post-training – pre-training) in symbolic quantity discrimination task efficiency – were not significantly different between the two training groups (two-sample t-test, t(50) = -.26, p = .80, Cohen's d = -.08, BF = .30). These results demonstrate that four weeks of number sense training improved symbolic quantity discrimination ability in both TD children and children with MLD.

Surprisingly, our sample of children with MLD did not perform poorly on symbolic quantity discrimination compared to TD children either before (t(50) = -.71, p = .48, Cohen's d = -.21, BF = .36) or after (t(50) = -1.01, p = .32, Cohen's d = -.30, BF = .44) training, with comparable variability between groups across time (coefficient of variation range: TD children: .23 – .31; children with MLD: .26 – .35; test of variance: Fs < 1.14, ps > .73). Post-hoc analysis revealed that children's performance on all measures of WJ-III math subtests were not significantly correlated with symbolic quantity discrimination before training in either group (ρs < .21, ps > .42, BFs < .65). These results suggest that there is a large variability in these children's ability to perform on basic numerical tasks and that mathematical difficulties may be present even in the absence of number sense deficits.

Association between intrinsic functional connectivity of the hippocampus and training-induced number sense learning

To test our central hypothesis that hippocampal functional circuits underpin number sense learning, we performed seed-to-whole-brain functional connectivity analyses using the left and right hippocampus ROIs derived from the Brainnetome (Fan et al., 2016). We first examined hippocampal connectivity patterns associated with number sense learning in the training group as a whole, and then followed up with analysis on TD children and children with MLD. As described in *Methods*, the analyses of associations between hippocampal connectivity and number sense learning controlled for pre-training symbolic quantity discrimination ability.

Training group. Functional connectivity of the left and right hippocampal ROIs with the left IPS before training was positively correlated with number sense learning in the training group (height threshold $p < .005$, cluster extent threshold $p < .05$) (**Figures 3A-B; Tables 6-7**). Similar results were observed at a more stringent height threshold ($p < .001$, uncorrected). A conjunction analysis of connectivity patterns across the left and right hippocampal ROIs confirmed a single overlapping target region in the left IPS, as identified by Juelich Histological Atlas, positively associated with number sense learning (**Figure 3C**). The left cuneus was identified as an overlapping target region negatively associated with number sense learning. Follow-up ROI-based correlation analyses were conducted in the training and control group, using the functional connectivity between both the left and right hippocampal regions and the left IPS region identified in the training group. Similar to results from whole-brain regression analysis, hippocampal functional connectivity with the left IPS predicted number sense learning in the training group (left hippocampus: $\rho = .42$, $p = .002$, BF = 18.10; right hippocampus: $\rho = .41$, $p = .003$, BF = 9.93; **Figures 4A-B**; **Table 8**). This association was not significant in the control group (left hippocampus: $\rho = -.20$, $p = .45$, BF = .56; right hippocampus: $\rho = -.14$, $p = .59$, BF = .53). A balanced fourfold cross-validation combined with linear regression (see *Methods*) further validated the robustness of findings in the training group. Functional connectivity of the left hippocampus [*r(pred,actual)* = .33, $p = .002$] and the right hippocampus [*r(pred,actual)* = .32,

$p = .002$] with the left IPS was predictive of gains in symbolic quantity discrimination task efficiency following training. This relationship was not significant in the control group for both the left and right hippocampal ROIs (ps > .50).

Direct comparisons of correlation coefficients between training and control groups revealed a significant difference in the relationship between functional connectivity with the left IPS and learning for the left ($Z = 2.15$, $p = .02$) and the right hippocampus ($Z = 1.90$, $p = .03$). Additional correlational analyses were conducted to examine the specificity of left-lateralized IPS functional connectivity patterns associated with learning. Here, the functional connectivity of the left and the right hippocampus with the right IPS (a contralateral region of the left IPS identified in the whole training group) did not significantly relate to number sense learning in the training or control group (ρs < .16, ps > .28, BFs < .56; **Table 8**).

Additional analysis confirmed that IPS regions identified from the current study overlap with the left IPS region identified from Neurosynth based meta-analysis, using the term "arithmetic" as defined in a previous study (Supekar et al., 2021) (**Figures 5A-B**), which indicates that the IPS region identified from our whole brain analysis converges with the region previously shown to be involved in math cognition. Finally, when using the IPS region identified from Neurosynth based meta-analysis, the association between hippocampal-left IPS circuits and learning in response to number sense training remains significant (left hippocampus-left IPS: $\rho = .35$, $p = .011$; right hippocampus-left IPS: $\rho = .30$, $p = .03$; **Figures 5C-D**).

Taken together, these results demonstrate that bilateral hippocampal functional connectivity with a common target in the left IPS is predictive of learning in response to a 4-week number sense training.

TD and MLD groups. We next examined whether hippocampal functional circuits predict number sense learning similarly or differently between children with and without MLD. We first separately conducted seed-to-whole-brain connectivity analyses for the left and right hippocampus in TD children and children with MLD. In TD children, functional connectivity of the left and right hippocampal ROIs with the left IPS predicted number sense learning (**Tables 6-7**), similar to the results in the whole training group. In a conjunction analysis of the connectivity patterns across the left and right hippocampal ROIs, TD children showed the left IPS as an overlapping region positively associated with learning, and the left cuneus as an overlapping region negatively associated with learning (**Figure 3D**), again replicating the results of the whole training group. In contrast to the distinctive pattern observed with the TD group whose functional connectivity between hippocampal regions and the left IPS was positively associated with learning, children with MLD showed no brain regions as targets from hippocampus-to-whole-brain connectivity positively associated with number sense learning, for either left or right hippocampal ROI (**Tables 6-7**). For the right hippocampus, its connectivity with the right cuneus was negatively associated with learning in children with MLD, a similar pattern observed in TD children though in the contralateral side of the cuneus. Considering the possibility that hippocampal–left IPS circuits were not detected at the whole brain level due to more heterogeneous sample in the MLD group, we next performed ROI-based correlation analyses for both groups.

Using the left IPS region identified in the whole training group as target ROI, we found that hippocampal functional connectivity with the left IPS is positively associated with number sense learning in the TD group (left hippocampus: $\rho = .43$, $p = .01$, BF = 3.98; right hippocampus: $\rho = .38$, $p = .03$, BF = 4.58; **Table 9**) as well as in children with MLD (left hippocampus: $\rho = .52$, $p = .03$, BF = 2.20; right hippocampus: $\rho = .52$, $p = .03$, BF = .93). In a cross-validation analysis (see *Methods*), functional connectivity of the left [*r(pred,actual*) = .26, $p = .02$] and right hippocampus [$r(pred,actual) = .37$, $p = .006$] with the left IPS was predictive of learning in TD children. In children with MLD, the relationship between functional connectivity and learning did not reach statistical significance at $p < .05$ in the left [$r(pred,actual) = 22$, $p = .06$] and right [$r(pred,actual) = .10$, $p = .15$] hippocampus.

Direct comparisons of correlation coefficients between TD and MLD groups revealed no significant difference in the relationship between hippocampal functional connectivity with the left IPS and learning gains (left hippocampus: $Z = .37$, $p = .36$; right hippocampus: $Z = .56$, $p = .29$). Finally, similar to the results from whole training group, functional connectivity of the left and right hippocampus with the right IPS did not significantly relate to learning in TD children or children with MLD (ρs < .33, *p*s > .19, BFs < .55; **Table 9**).

To further address whether the relation between hippocampal-parietal functional connectivity and number sense learning varies as a function of individual differences in math ability, we additionally used a dimensional approach. In a multiple regression model, number sense learning (gains in symbolic quantity discrimination task efficiency) was entered as dependent variable, hippocampal-parietal connectivity (left hippocampus – left IPS or right hippocampus – left IPS link), math ability (WJ-III Math Fluency), and interaction between hippocampal-parietal connectivity and math ability were entered as independent variables, and pre-training symbolic number comparison efficiency was entered as a covariate. We found a significant main effect of hippocampal-parietal connectivity (left hippocampus – left IPS: $b = .36$, $se = .12$, $t = 3.13$, $p = .003$; right hippocampus – left IPS: $b = .32$, $se = .12$, $t = 2.57$, $p = .01$) but no significant main effect of math ability or interaction between hippocampal-parietal connectivity and math ability (*t*s < .81, *p*s > .42) on number sense learning.

In summary, in TD children, both left and right hippocampal functional connectivity with the left IPS predicted training-induced number sense learning, similar to the whole training group. Although hippocampal–left IPS circuits were not detected at the whole brain level, children with MLD showed hippocampal–left IPS functional connectivity associated with number sense learning in ROI-based analysis, which was relatively weaker but not significantly different from TD children. Finally, additional analysis using math ability as a continuous variable confirmed that the relation between hippocampal-parietal connectivity and learning did not significantly vary between individuals with different levels of math ability.

Association between hippocampal–parietal white matter pathways and training-induced number sense learning

To determine whether structural integrity plays a similar role in learning as functional circuitry, we examined the relation between pre-training hippocampal–parietal white matter connectivity and number sense learning. Using probabilistic tractography of HARDI data, we identified long-range anatomical connections between the hippocampus and IPS, identified from functional connectivity analysis, averaged across all children (**Figure 6**; see *Methods* for details) and assessed the relation between hippocampal–parietal structural connectivity and number sense learning. In contrast to the functional connectivity results, however, we did not observe evidence for structural connectivity of the hippocampus with IPS associated with training-related gains in symbolic quantity discrimination task ($|\rho|$s < .21, ps > .18, BFs < .62; **Table 8**). Further, there were no significant associations between structural connectivity between hippocampus and IPS and number sense learning in TD children and children with MLD ($|\rho|$s < .39, ps > .16, BFs < 1.34; **Table 9**).

To further address potential contribution of structural connectivity measures to learning, we conducted multiple regression analysis to determine whether functional and structural connectivity measures together predicted number sense learning better than each measure alone. Specifically, we examined whether the full model (Model 4) including all hippocampal–IPS functional and structural connectivity measures as predictors, compared to including functional or structural connectivity measures alone (Model 2 and Model 3, respectively), better predict number sense learning (**Table 10**). Here we found that the full model (Model 4) explained most variance in number sense learning (adjusted $R^2 = .54$, $F(9,33) = 6.55$, $p < .001$), significantly better than the model including structural connectivity measures alone (Model 3; $\Delta R^2 = .33$, $p < .001$, BF > 100). Critically, there was insufficient evidence (.33 < BF < 3) that the full model including both functional and structural connectivity measures (Model 4) explain additional variance in learning, compared to the model including functional connectivity measures alone (Model 2; $\Delta R^2 = .14$, $p = .026$, BF = 2.26). Thus, we did not observe evidence that structural connectivity measures jointly predict number sense learning over and above functional connectivity measures.

Taken together, these results suggest that the integrity of white matter pathways between the hippocampus and IPS in early childhood is not predictive of number sense learning in the current study. In addition, we observed evidence for joint associations between hippocampal–IPS functional circuits and number sense learning, independent of the underlying structural connectivity.

Reverse meta-analysis of associations between hippocampal–parietal circuits and cognitive functions

To further determine the functional role of hippocampal–IPS circuits identified in the current study, we conducted a reverse meta-analysis across 14,371 published fMRI studies up to July 2018 from NeuroSynth (Yarkoni et al., 2011) database in relation to 89 cognitive atlas terms CogAt (Poldrack et al., 2011). To perform reverse meta-analysis relating a cognitive function with a specific circuit, we computed the probability that a term associated with a cognitive function is mentioned in a study, given that the study jointly reports activations in the left or right hippocampus and left or right IPS. We considered sufficient evidence for an association if its probability is amongst the 5% most probable term associations for all analyzed circuits (see *Methods* for details). This meta-analysis tool allowed us to synthesize a wealth of findings from previous research and generalize the findings of hippocampal–parietal circuits across various tasks and analysis approaches (Muller et al., 2018).

Our results from reverse meta-analysis show that co-activations of both the left and right hippocampus and IPS are significantly associated with the term *learning* as well as related terms, *encoding*, *memory,* and *retrieval* (**Figure 7**). The term *recognition* was associated with co-activations of the left hippocampus and left IPS and those of bilateral hippocampus and right IPS. Two terms, *attention* and *working memory,* were associated with co-activations of the right hippocampus and bilateral IPS. The term *emotion* was associated with co-activations of the right hippocampus and right IPS. Finally, the term *perception* was associated with co-activations of the left hippocampus and left IPS. Notably, no other cognitive atlas terms were significantly associated with hippocampal–parietal functional circuits. These meta-analytic findings from a large set of fMRI studies expand on findings from our training study and provide converging evidence for a strong association between hippocampal–parietal functional circuitry and learning and related functions.

Discussion

The current study examined brain circuit mechanisms of learning in response to an integrative number sense training during an important developmental period for foundational cognitive skill acquisition. Our results reveal that number sense training significantly improves symbolic quantity discrimination ability in both TD children and children with MLD, and that hippocampal–

left IPS functional circuits predict number sense training gains. Our findings provide important insights into brain-based biomarkers for early identification of individual differences in acquisition of number sense and inform interventions targeting individual needs (Hale et al., 2010).

We show that our integrative number sense training is effective across children with a wide range of abilities. Our results build on previous training studies that enhance understanding of numerical magnitudes in children (Dyson, Jordan, & Glutting, 2013; Kucian et al., 2011; Wilson et al., 2006). Our study maximized effectiveness of training by uniquely combining computerized games with tutoring activities using physical manipulatives and provide new insights into the development of effective ‘hybrid’ interventions across children with different backgrounds. More generally, individualized training programs designed to enhance integration of symbolic and non-symbolic representations of quantity may have the potential to build strong foundations for mathematical learning across all children.

It is noteworthy that prior to training, we did not observe poor number sense in our sample of children with MLD, who performed significantly worse than TD children on assessments of arithmetic problem solving and mathematical reasoning, similar to previous observation that difficulties in math problem solving may be present even in the absence of number sense deficits (Peters, Op de Beeck, & De Smedt, 2020). As MLD is considered a heterogeneous disorder with multiple cognitive deficits (Fias, Menon, & Szucs, 2013; Kaufmann et al., 2013), further studies that employ assessments in various cognitive domains may help determine multidimensional neurocognitive deficits in MLD. In addition, development of classification of subtypes of MLD will be an important avenue for future research.

Our next goal was to investigate whether the integrity of hippocampal–parietal circuits predicts individual differences in number sense training gains in children. We identified an intrinsic functional circuit that links the hippocampus, a hub for learning and memory, with a parietal region consistently implicated in numerical quantity representation. Our finding converges on previous studies demonstrating the key functional role of the hippocampus in the development of arithmetic skills in children (Menon, 2016; Menon & Chang, 2021) and learning and memory more broadly (Zeithamova & Bowman, 2020). Notably, this contribution occurred even though our number sense training did not require rote memorization of facts, and is consistent with emerging evidence for a role of the hippocampus that extends beyond explicit memory and learning (Degonda et al., 2005; Olsen et al., 2012).

Our finding is also consistent with previous evidence indicating the role of IPS in representation of quantities (Cohen Kadosh, Cohen Kadosh, Kaas, Henik, & Goebel, 2007; Piazza & Eger, 2016). Remarkably, functional connectivity of the left and right hippocampus identified a single region in the left IPS that predicted learning. Previous studies have found that compared to its right hemisphere homolog, the left IPS is particularly important for symbolic number processing (Ansari, 2007; Bugden, Price, McLean, & Ansari, 2012; Piazza, Pinel, Le Bihan, & Dehaene, 2007; Sokolowski, Fias, Mousa, & Ansari, 2017) and that with age and increased proficiency in numerical skills, there is an increase in left IPS activity (Ansari, 2008; Bugden, DeWind, & Brannon, 2016; Emerson & Cantlon, 2015; Rivera, Reiss, Eckert, & Menon, 2005). In this context, it is possible that our finding of left-lateralized IPS response may be reflective of increased proficiency for symbolic numbers.

In addition to the left IPS as a converging target region for the left and right hippocampal functional circuits positively associated with number sense learning, the left cuneus, implicated in low-level visual processing (Vanni, Tanskanen, Seppa, Uutela, & Hari, 2001), was identified as a target region negatively associated with number sense training gains. This finding suggests that children's number sense learning likely relies on enhanced semantic representation of quantity, rather than visual perception of numbers. In fact, greater hippocampal functional connectivity with a visual region (cuneus) predicted poor learning. No other brain regions were identified as overlapping target regions across the left and right hippocampal functional connectivity associated with learning. Taken together, our findings demonstrate a key role for hippocampal–parietal circuits in children's number sense learning.

Our whole brain analysis of subgroups of children revealed that TD children recapitulate the hippocampal–left IPS functional circuit-related learning as seen in the combined group. While no significant target regions were detected in hippocampal connectivity positively associated with learning in children with MLD possibly due to modest sample size in this group, additional analysis confirmed that the association between hippocampal–left IPS circuits and learning was similar in the TD and MLD groups, consistent with our observation of comparable training gains in the two groups. Thus, our findings identify hippocampal–left IPS functional circuit as a novel locus of learning that supports acquisition of fundamental building blocks of numerical proficiency across all children, including those with learning disabilities. Future studies with a larger sample of children with MLD may further clarify heterogeneous profiles of learning-related hippocampal circuits.

Our probabilistic tractography of HARDI data indicates the presence of long-range anatomical connections between the hippocampus and the parietal cortex in 7-10 year old children, similar to the observation in younger children (Ngo et al., 2017). In contrast to findings from functional circuit analysis, however, structural connectivity between the hippocampus and IPS did not relate to individual differences in learning. In addition, while hippocampal–parietal functional connectivity measures jointly predicted learning over and above measures of structural integrity, we did not find evidence that structural connectivity measures jointly predict learning over and above functional connectivity measures. Thus, although hippocampal–parietal white matter tracts appear to be formed by early childhood, how they contribute to number sense learning remains unresolved. Crucially, our study provides evidence for emergent functional properties of hippocampal–parietal circuits as significant and independent neural predictors of number sense learning, which may inform early identification of individual differences in response to intervention. Future studies employing larger data sets with an active control group may help identify generalizable predictive features of learning and further determine specific mechanisms underlying acquisition of foundational cognitive skills.

Finally, our findings converge on results from a reverse meta-analysis in which we examined the role of hippocampal–parietal functional links identified in the present study. Across 14,371 fMRI studies and 89 cognitive atlas terms, our analysis revealed a significant association between bilateral hippocampal–IPS functional circuits and the term *learning*, along with related terms *memory, encoding,* and *retrieval*. Interactions between the hippocampus and neocortex are known to be crucial for memory formation (McClelland, McNaughton, & O'Reilly, 1995; Tse et al., 2007) and hippocampal connectivity with parietal and frontal cortical regions have been shown to be associated with longitudinal gains in memory retrieval fluency in children (Qin et al., 2014). Taken together, our findings identify specific hippocampal–neocortical functional circuitries that may contribute to learning and memory consolidation.

While involvement of the hippocampus in learning and memory is well known, the specific role of hippocampal–IPS functional circuits has been less clear, as research on the role of parietal cortex in memory has emphasized its angular gyrus subdivision in episodic memory (Sestieri, Shulman, & Corbetta, 2017). The angular gyrus, as part of the default mode network, is crucial for generating integrated representation of information retrieved from episodic memory (Binder

& Desai, 2011). In contrast, the IPS, as part of the dorsal attention network, is crucial for representing and manipulating visuospatial perceptual information (Uddin et al., 2010). Consistent with this view, we found that IPS-relevant terms, specifically *perception, attention,* and *working memory*, are associated with hippocampus–IPS circuits. In the context of number sense learning, we propose that these functions, together with learning and memory, support the formation of semantic associations between quantities presented in non-symbolic and symbolic formats. Thus, findings from a reverse meta-analysis of a large corpus of fMRI studies and our training study suggest that hippocampal–IPS circuits constitute a distinct canonical circuit for integrating and manipulating mnemonic and visuospatial information that plays a foundational role in children's cognitive skill acquisition. More broadly, interventions designed to engage these brain circuits, such as integrative number sense training in the current study, may effectively promote learning across various cognitive domains.

In summary, the current study demonstrates that core learning and memory functional circuits anchored in the hippocampus play an important role in learning number sense, a fundamental building block of mathematical skill acquisition. Notably, the left IPS, implicated in numerical proficiency, was a convergence zone for the left and right hippocampal functional circuits that predict individual differences in number sense learning. Our study provides foundational knowledge about brain circuit mechanisms that propel learning in all children and delineates a robust target for effective interventions and monitoring response to cognitive training.

References

Andersson, J. L. R., & Sotiropoulos, S. N. (2016). An integrated approach to correction for off-resonance effects and subject movement in diffusion MR imaging. *NeuroImage, 125*, 1063-1078. doi:10.1016/j.neuroimage.2015.10.019

Ansari, D. (2007). Does the parietal cortex distinguish between "10," "ten," and ten dots? *Neuron, 53*(2), 165-167. doi:10.1016/j.neuron.2007.01.001

Ansari, D. (2008). Effects of development and enculturation on number representation in the brain. *Nature reviews neuroscience, 9*, 278-291. doi:10.1038/nrn2334

Avants, B. B., Schoenemann, P. T., & Gee, J. C. (2006). Lagrangian frame diffeomorphic image registration: Morphometric comparison of human and chimpanzee cortex. *Med Image Anal, 10*(3), 397-412. doi:10.1016/j.media.2005.03.005

Barnett, A. G., van der Pols, J. C., & Dobson, A. J. (2005). Regression to the mean: What it is and how to deal with it. *International Journal of Epidemiology, 34*, 215–220.

Behrens, T. E., Berg, H. J., Jbabdi, S., Rushworth, M. F., & Woolrich, M. W. (2007). Probabilistic diffusion tractography with multiple fibre orientations: What can we gain? *NeuroImage, 34*(1), 144-155. doi:10.1016/j.neuroimage.2006.09.018

Binder, J. R., & Desai, R. H. (2011). The neurobiology of semantic memory. *Trends Cogn Sci, 15*(11), 527-536. doi:10.1016/j.tics.2011.10.001

Blair, K. P. (2013a). *Feedback in Critter Corral: The effectiveness of implication versus corrective feedback in a math learning game*. Paper presented at the 2013 Early Education and Technology for Children Conference, Salt Lake City, UT.

Blair, K. P. (2013b). *Learning in Critter Corral: Evaluating three kinds of feedback in a preschool math game.* Paper presented at the Proceedings of the Interaction Design and Children 2013 Conference, New York: ACM.

Bugden, S., DeWind, N. K., & Brannon, E. M. (2016). Using cognitive training studies to unravel the mechanisms by which the approximate number system supports symbolic math ability. *Curr Opin Behav Sci, 10*, 73-80. doi:10.1016/j.cobeha.2016.05.002

Bugden, S., Price, G. R., McLean, D. A., & Ansari, D. (2012). The role of the left intraparietal sulcus in the relationship between symbolic number processing and children's arithmetic competence. *Dev Cogn Neurosci, 2*(4), 448-457. doi:10.1016/j.dcn.2012.04.001

Butterworth, B., & Walsh, V. (2011). Neural basis of mathematical cognition. *Current biology : CB, 21*, R618-621. doi:10.1016/j.cub.2011.07.005

Carter, C. S., Lesh, T. A., & Barch, D. M. (2016). Thresholds, power, and sample sizes in clinical neuroimaging. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging, 1*, 99–100.

Chang, H., Rosenberg-Lee, M., Qin, S., & Menon, V. (2019). Faster learners transfer their knowledge better: Behavioral, mnemonic, and neural mechanisms of individual differences in children's learning. *Dev Cogn Neurosci, 40*, 100719. doi:10.1016/j.dcn.2019.100719

Cho, S., Metcalfe, A. W. S., Young, C. B., Ryali, S., Geary, D. C., & Menon, V. (2012). Hippocampal–Prefrontal Engagement and Dynamic Causal Interactions in the Maturation of Children's Fact Retrieval. *Journal of Cognitive Neuroscience, 24*, 1849-1866. doi:10.1162/jocn_a_00246

Cho, S., Ryali, S., Geary, D. C., & Menon, V. (2011). How does a child solve 7+8? Decoding brain activity patterns associated with counting and retrieval strategies. *Developmental science, 14*, 989-1001. doi:10.1111/j.1467-7687.2011.01055.x

Cohen, J. R., Asarnow, R. F., Sabb, F. W., Bilder, R. M., Bookheimer, S. Y., Knowlton, B. J., & Poldrack, R. A. (2010). Decoding developmental differences and individual variability in response inhibition through predictive analyses across individuals. *Front Hum Neurosci, 4*, 47. doi:10.3389/fnhum.2010.00047

Cohen Kadosh, R., Cohen Kadosh, K., Kaas, A., Henik, A., & Goebel, R. (2007). Notation-dependent and -independent representations of numbers in the parietal lobes. *Neuron, 53*(2), 307-314. doi:10.1016/j.neuron.2006.12.025

De Smedt, B., & Gilmore, C. K. (2011). Defective number module or impaired access? Numerical magnitude processing in first graders with mathematical difficulties. *Journal of experimental child psychology, 108*(2), 278-292. doi:10.1016/j.jecp.2010.09.003

Degonda, N., Mondadori, C. R., Bosshardt, S., Schmidt, C. F., Boesiger, P., Nitsch, R. M., . . . Henke, K. (2005). Implicit associative learning engages the hippocampus and interacts with explicit associative learning. *Neuron, 46*(3), 505-520. doi:10.1016/j.neuron.2005.02.030

Desikan, R., Ségonne, F., Fischl, B., Quinn, B., Dickerson, B., Blacker, D., . . . Killiany, R. (2006). An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *NeuroImage, 31*, 968-980.

Donahue, C. J., Sotiropoulos, S. N., Jbabdi, S., Hernandez-Fernandez, M., Behrens, T. E., Dyrby, T. B., . . . Glasser, M. F. (2016). Using Diffusion Tractography to Predict Cortical Connection Strength and Distance: A Quantitative Comparison with Tracers in the Monkey. *The Journal of Neuroscience, 36*(25), 6758-6770. doi:10.1523/jneurosci.0493-16.2016

Dyson, N. I., Jordan, N. C., & Glutting, J. (2013). A Number Sense Intervention for Low-Income Kindergartners at Risk for Mathematics Difficulties. . *Journal of Learning Disabilities, 46* 166–181.

Eichenbaum, H. (2004). Hippocampus: Cognitive processes and neural representations that underlie declarative memory. *Neuron, 44*(1), 109-120. doi:DOI 10.1016/j.neuron.2004.08.028

Eickhoff, S. B., Stephan, K. E., Mohlberg, H., Grefkes, C., Fink, G. R., Amunts, K., & Zilles, K. (2005). A new SPM toolbox for combining probabilistic cytoarchitectonic maps and functional imaging data. *NeuroImage, 25*, 1325-1335.

Emerson, R. W., & Cantlon, J. F. (2015). Continuity and change in children's longitudinal neural responses to numbers. *Developmental Science, 18*, 314-326. doi:10.1111/desc.12215

Fan, L., Li, H., Zhuo, J., Zhang, Y., Wang, J., Chen, L., . . . Jiang, T. (2016). The Human Brainnetome Atlas: A New Brain Atlas Based on Connectional Architecture. *Cereb Cortex, 26*(8), 3508-3526. doi:10.1093/cercor/bhw157

Fias, W., Menon, V., & Szucs, D. (2013). Multiple components of developmental dyscalculia. *Trends in Neuroscience and Education, 2*(2), 43-47. doi:10.1016/j.tine.2013.06.006

Forman, S. D., Cohen, J. D., Fitzgerald, M., Eddy, W. F., Mintun, M. A., & Noll, D. C. (1995). Improved assessment of significant activation in functional magnetic resonance imaging (fMRI): use of a cluster-size threshold. *Magn. Reson. Med., 33*, 636-647.

Fuchs, L. S., Powell, S. R., Seethaler, P. M., Cirino, P. T., Fletcher, J. M., Fuchs, D., . . . Zumeta, R. O. (2009). Remediating Number Combination and Word Problem Deficits Among Students With Mathematics Difficulties: A Randomized Control Trial. *Journal of Educational Psychology, 101*, 561–576.

Glover, G. H., & Lai, S. (1998). Self-navigated spiral fMRI: Interleaved versus single-shot. *Magnetic Resonance in Medicine, 39*, 361–368. .

Greicius, M. D., Krasnow, B., Reiss, A. L., & Menon, V. (2003). Functional connectivity in the resting brain: a network analysis of the default mode hypothesis. *Proc.Natl.Acad.Sci.U.S A, 100*, 253-258. doi:10.1073/pnas.0135058100

Hale, J., Alfonso, V., Berninger, V., Bracken, B., Christo, C., Clark, E., . . . Yalof, J. (2010). Critical issues in response-to-intervention, comprehensive evaluation, and specific learning disabilities identification and intervention: An expert white paper consensus. *Learning Disability Quarterly, 33*, 223-236.

Iovene, V., & Wassermann, D. (2020). *Probabilistic Programming in neurolang: Bridging the Gap Between Cognitive Science and Statistical Modeling*. Paper presented at the Organization for Human Brain Mapping, Remote due to Covid-19, France.

Iuculano, T., Rosenberg-Lee, M., Richardson, J., Tenison, C., Fuchs, L., Supekar, K., & Menon, V. (2015). Cognitive tutoring induces widespread neuroplasticity and remediates brain function in children with mathematical learning disabilities. *Nature Communications, 6*, 8453. doi:10.1038/ncomms9453

Iuculano, T., Rosenberg-Lee, M., Supekar, K., Lynch, C. J., Khouzam, A., Phillips, J., . . . Menon, V. (2014). Brain organization underlying superior mathematical abilities in children with autism. *Biol Psychiatry, 75*(3), 223-230. doi:10.1016/j.biopsych.2013.06.018

Jolles, D., Ashkenazi, S., Kochalka, J., Evans, T., Richardson, J., Rosenberg-Lee, M., . . . Menon, V. (2016). Parietal hyper-connectivity, aberrant brain organization, and circuit-based biomarkers in children with mathematical disabilities. *Developmental science, 19*, 613-631. doi:10.1111/desc.12399

Jordan, N. C., Kaplan, D., Ramineni, C., & Locuniak, M. N. (2009). Early math matters: kindergarten number competence and later mathematics outcomes. *Dev Psychol, 45*(3), 850-867. doi:10.1037/a0014939

Kaufmann, L., Mazzocco, M. M., Dowker, A., von Aster, M., Gobel, S. M., Grabner, R. H., . . . Nuerk, H. C. (2013). Dyscalculia from a developmental and differential perspective. *Front Psychol, 4*, 516. doi:10.3389/fpsyg.2013.00516

Kaufmann, L., & von Aster, M. (2012). The diagnosis and management of dyscalculia. *Dtsch Arztebl Int, 109*(45), 767-777; quiz 778. doi:10.3238/arztebl.2012.0767

Keysers, C., Gazzola, V., & Wagenmakers, E. J. (2020). Using Bayes factor hypothesis testing in neuroscience to establish evidence of absence. *Nat Neurosci, 23*(7), 788-799. doi:10.1038/s41593-020-0660-4

Kim, D.-H., Adalsteinsson, E., Glover, G. H., & Spielman, D. M. (2002). Regularized higher-order in vivo shimming. *Magnetic Resonance in Medicine, 48*, 715–722.

Kucian, K., Grond, U., Rotzer, S., Henzi, B., Schonmann, C., Plangger, F., . . . von Aster, M. (2011). Mental number line training in children with developmental dyscalculia. *NeuroImage, 57*(3), 782-795. doi:10.1016/j.neuroimage.2011.01.070

Landerl, K., Göbel, S. M., & Moll, K. (2013). Core deficit and individual manifestations of developmental dyscalculia (DD): The role of comorbidity. *Trends in Neuroscience and Education, 2*(2), 38-42. doi:10.1016/j.tine.2013.06.002

McClelland, J. L., McNaughton, B. L., & O'Reilly, R. C. (1995). Why there are complementary learning systems in the hippocampus and neocortex: insights from the successes and failures of connectionist models of learning and memory. *Psychol Rev, 102*(3), 419-457. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/7624455

Menon, V. (2016). Memory and cognitive control circuits in mathematical cognition and learning. *Prog Brain Res, 227*, 159-186. doi:10.1016/bs.pbr.2016.04.026

Menon, V., & Chang, H. (2021). Emerging neurodevelopmental perspectives on mathematical learning *Developmental Review, 60*.

Muller, V. I., Cieslik, E. C., Laird, A. R., Fox, P. T., Radua, J., Mataix-Cols, D., . . . Eickhoff, S. B. (2018). Ten simple rules for neuroimaging meta-analysis. *Neurosci Biobehav Rev, 84*, 151-161. doi:10.1016/j.neubiorev.2017.11.012

National Mathematics Advisory Panel. (2008). *Foundations for Success: The Final Report of the National Mathematics Advisory Panel*. Retrieved from Washington, DC:

Ngo, C. T., Alm, K. H., Metoki, A., Hampton, W., Riggins, T., Newcombe, N. S., & Olson, I. R. (2017). White matter structural connectivity and episodic memory in early childhood. *Dev Cogn Neurosci, 28*, 41-53. doi:10.1016/j.dcn.2017.11.001

Nichols, T., & Hayasaka, S. (2003). Controlling the familywise error rate in functional neuroimaging: a comparative review. *Stat. Methods Med. Res., 12*, 419-446.

Olsen, R. K., Moses, S. N., Riggs, L., & Ryan, J. D. (2012). The hippocampus supports multiple cognitive processes through relational binding and comparison. *Front Hum Neurosci, 6*, 146. doi:10.3389/fnhum.2012.00146

Peters, L., Op de Beeck, H., & De Smedt, B. (2020). Cognitive correlates of dyslexia, dyscalculia and comorbid dyslexia/dyscalculia: Effects of numerical magnitude processing and phonological processing. *Research in Developmental Disabilities, 107*(March), 103806-103806. doi:10.1016/j.ridd.2020.103806

Piazza, M., & Eger, E. (2016). Neural foundations and functional specificity of number representations. *Neuropsychologia, 83*, 257-273. doi:10.1016/j.neuropsychologia.2015.09.025

Piazza, M., Pinel, P., Le Bihan, D., & Dehaene, S. (2007). A magnitude code common to numerosities and number symbols in human intraparietal cortex. *Neuron, 53*(2), 293-305. doi:10.1016/j.neuron.2006.11.022

Pocock, S. J., Assmann, S. E., Enos, L. E., & Kasten, L. E. (2002). Subgroup analysis, covariate adjustment and baseline comparisons in clinical trial reporting: Current practice and problems. *Statistics in Medicine,, 21*, 2917–2930.

Poldrack, R. A., Kittur, A., Kalar, D., Miller, E., Seppa, C., Gil, Y., . . . Bilder, R. M. (2011). The Cognitive Atlas: Toward a Knowledge Foundation for Cognitive Neuroscience. *Frontiers in Neuroinformatics, 5*, 1-11. doi:10.3389/fninf.2011.00017

Qin, S., Cho, S., Chen, T., Rosenberg-Lee, M., Geary, D. C., & Menon, V. (2014). Hippocampal-neocortical functional reorganization underlies children's cognitive development. *Nature Neuroscience*. doi:10.1038/nn.3788

Richland, L. E., Kornell, N., & Kao, L. S. (2009). The pretesting effect: do unsuccessful retrieval attempts enhance learning? *J Exp Psychol Appl, 15*(3), 243-257. doi:10.1037/a0016496

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 (New York, N.Y. : 1991), 15*, 1779-1790. doi:10.1093/cercor/bhi055

Roediger, H. L., Agarwal, P. K., McDaniel, M. A., & McDermott, K. B. (2011). Test-enhanced learning in the classroom: long-term improvements from quizzing. *J Exp Psychol Appl, 17*(4), 382-395. doi:10.1037/a0026252

Rosenberg-Lee, M., Ashkenazi, S., Chen, T., Young, C. B., Geary, D. C., & Menon, V. (2015). Brain hyper-connectivity and operation-specific deficits during arithmetic problem solving in children with developmental dyscalculia. *Developmental Science, 18*, 351-372. doi:10.1111/desc.12216

Rosenberg-Lee, M., Barth, M., & Menon, V. (2011). What difference does a year of schooling make? Maturation of brain response and connectivity between 2nd and 3rd grades during arithmetic problem solving. *NeuroImage, 57*(3), 796-808. doi:10.1016/j.neuroimage.2011.05.013

Rosenberg-Lee, M., Iuculano, T., Bae, S. R., Richardson, J., Qin, S., Jolles, D., & Menon, V. (2018). Short-term cognitive training recapitulates hippocampal functional changes associated with one year of longitudinal skill development. *Trends in Neuroscience and Education*. doi:10.1016/j.tine.2017.12.001

Rousselle, L., & Noël, M.-P. (2007). Basic numerical skills in children with mathematics learning disabilities: a comparison of symbolic vs non-symbolic number magnitude processing. *Cognition, 102*, 361-395. doi:10.1016/j.cognition.2006.01.005

Sestieri, C., Shulman, G. L., & Corbetta, M. (2017). The contribution of the human posterior parietal cortex to episodic memory. *Nat Rev Neurosci, 18*(3), 183-192. doi:10.1038/nrn.2017.6

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. doi:10.1016/j.neuroimage.2016.10.028

Supekar, K., Chang, H., Mistry, P. K., Iuculano, T., & Menon, V. (2021). Cognitive training-induced modular reorganization of hippocampal-cortical circuits drives learning gains and efficient strategies. *Communications Biology, 4*.

Supekar, K., Swigart, A. G., Tenison, C., Jolles, D. D., Rosenberg-lee, M., Fuchs, L., & Menon, V. (2013). Neural predictors of individual differences in response to math tutoring in primary-grade school children. *Proc Natl Acad Sci, 110*, 8230-8235.

Thomas, C., Ye, F. Q., Irfanoglu, M. O., Modi, P., Saleem, K. S., Leopold, D. A., & Pierpaoli, C. (2014). Anatomical accuracy of brain connections derived from diffusion MRI tractography is inherently limited. *Proc Natl Acad Sci U S A, 111*(46), 16574-16579. doi:10.1073/pnas.1405672111

Thompson, D. D., Lingsma, H. F., Whiteley, W. N., Murray, G. D., & Steyerberg, E. W. (2015). Covariate adjustment had similar benefits in small and large randomized controlled trials. *Journal of Clinical Epidemiology, 68*, 1068–1075.

Townsend, J. T., & Ashby, F. G. (1978). *Methods of modeling capacity in simple processing systems* (Vol. 3). Erlbaum, Hillsdale, N.J.: Cognitive Theory.

Tse, D., Langston, R. F., Kakeyama, M., Bethus, I., Spooner, P. A., Wood, E. R., . . . Morris, R. G. M. (2007). Schemas and Memory Consolidation. *Science, 316*, 76–82.

Tuch, D. S., Reese, T. G., Wiegell, M. R., Makris, N., Belliveau, J. W., & Wedeen, V. J. (2002). High angular resolution diffusion imaging reveals intravoxel white matter fiber heterogeneity. *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine, 48*, 577-582.

Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., . . . Joliot, M. (2002). Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *NeuroImage, 15* 273-289.

Uddin, L. Q., Supekar, K., Amin, H., Rykhlevskaia, E., Nguyen, D. A., Greicius, M. D., & Menon, V. (2010). Dissociable connectivity within human angular gyrus and intraparietal sulcus: evidence from functional and structural connectivity. *Cereb Cortex, 20*(11), 2636-2646. doi:10.1093/cercor/bhq011

Vanni, S., Tanskanen, T., Seppa, M., Uutela, K., & Hari, R. (2001). Coinciding early activation of the human primary visual cortex and anteromedial cuneus. *Proc Natl Acad Sci U S A, 98*(5), 2776-2780. doi:10.1073/pnas.041600898

Ward, B. D. (2000). Simultaneous Inference for FMRI Data. In *AFNI 3dDeconvolve Documentation*. Milwaukee, WI: Medical College of Wisconsin.

Wechsler, D. (1999). *The Wechsler Abbreviated Scale of Intelligence*. San Antonio, TX: The Psychological Corp.

Wilson, A. J., Revkin, S. K., Cohen, D., Cohen, L., & Dehaene, S. (2006). An open trial assessment of "the number race", an adaptive computer game for remediation of dyscalculia. *Behavioral and Brain Functions, 2*, 20. doi:10.1186/1744-9081-2-20

Woodcock, R. W., McGrew, K. S., & Mather, N. (2001). *Woodcock–Johnson III Tests of Achievement.* Istasca, IL: Riverside.

Yarkoni, T., Poldrack, R. a., Nichols, T. E., Van Essen, D. C., & Wager, T. D. (2011). Large-scale automated synthesis of human functional neuroimaging data. *Nature methods, 8*, 665-670. doi:10.1038/nmeth.1635

Zeithamova, D., & Bowman, C. R. (2020). Generalization and the hippocampus: More than one story? *Neurobiology of Learning and Memory, 175*, 107317 Contents.

Figures and Tables

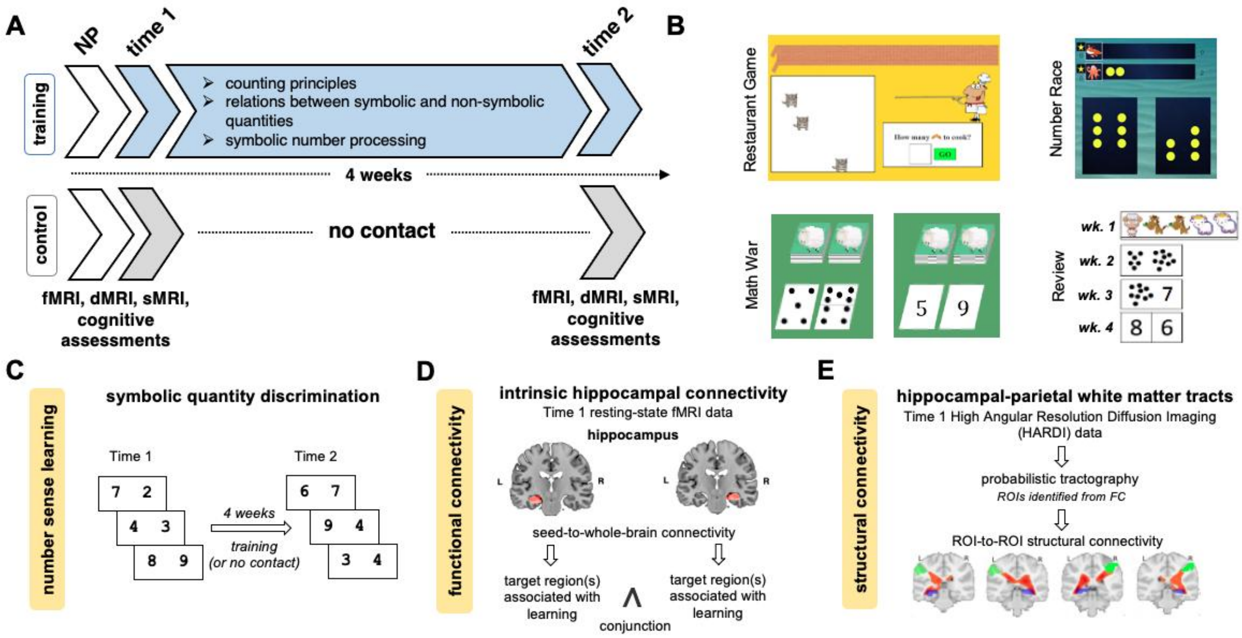

Figure 1. Overview of study design, sample training materials, and schematic illustration of analysis approach.

A. ***Overview of study design***. The study involved multiple sessions: pre-training demographic and neuropsychological (NP) assessments; pre-training (time 1) cognitive assessments and brain imaging, including task-related and resting-state functional MRI (fMRI), diffusion MRI (dMRI) using a High Angular Resolution Diffusion Imaging sequence, and high-resolution structural MRI (sMRI); number sense training (in the training group; see *Methods* for details) or no contact (in the control group); and post-training (time 2) brain imaging and cognitive assessments identical to pre-training. Children in the training group engaged in progressive learning activities across four weeks to strengthen their understanding of the relations between symbolic and non-symbolic representations of quantity. **B**. ***Sample training materials***. Across 4 weeks of one-on-one tutoring sessions, children in the training group completed a variety of activities with a tutor (see *Methods* for details). **C. *Number sense learning***. Children's number sense learning was measured by changes in efficiency in symbolic quantity discrimination task from time 1 to time 2 in response to 4 weeks of number sense training (or no contact). **D.** ***Functional connectivity***. Using Time 1 resting-state fMRI data, hippocampal seed-to-whole-brain connectivity analysis was performed to assess intrinsic functional connectivity of the hippocampus predictive of number sense learning. A conjunction analysis was performed between left and right hippocampal functional connectivity patterns to identify overlapping target regions associated with learning. Region of interest (ROI) based analysis was used to examine the association between functional connectivity and learning in training and control groups. **E.** ***Structural connectivity***. Using High Angular Resolution Diffusion Imaging data, probabilistic tractography was performed using ROIs identified from functional connectivity (FC) analysis. Structural connectivity strengths between the ROIs were estimated for each subject to test the associations with number sense learning. ROI-based analysis was performed to examine the relation between structural connectivity and learning in training and control groups.

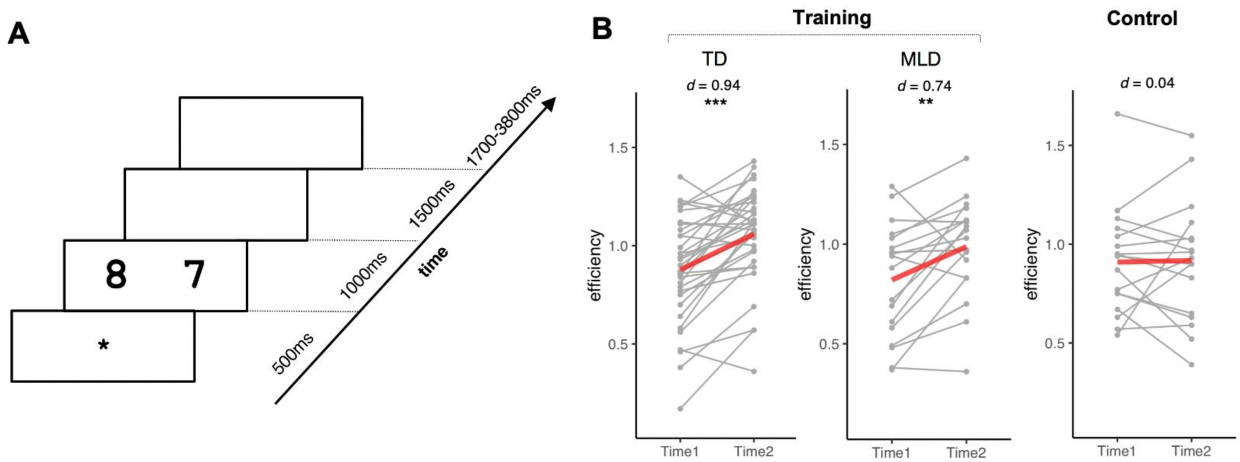

Figure 2. Symbolic quantity discrimination task design and behavioral results. **A**. ***Task design***. Before and after training (or no contact), children completed one run of symbolic quantity discrimination task in the fMRI scanner. Participants first saw a fixation cross, followed by horizontal presentation of two quantities in Arabic numbers. Participants were instructed to press the left button if the left side had a larger quantity and the right button if the right side had a larger quantity. Upon the button press, a blank screen was presented to fill up the response phase, followed by a jittered intertrial interval. The duration of each phase in seconds is reported in parenthesis. **B**. ***Behavioral results***. Four weeks of training improved performance on symbolic quantity discrimination task in both training groups of typically developing (TD) children and children with mathematical learning difficulties (MLD) but not in no-contact control group. ** *p* < .01; *** *p* < .001.

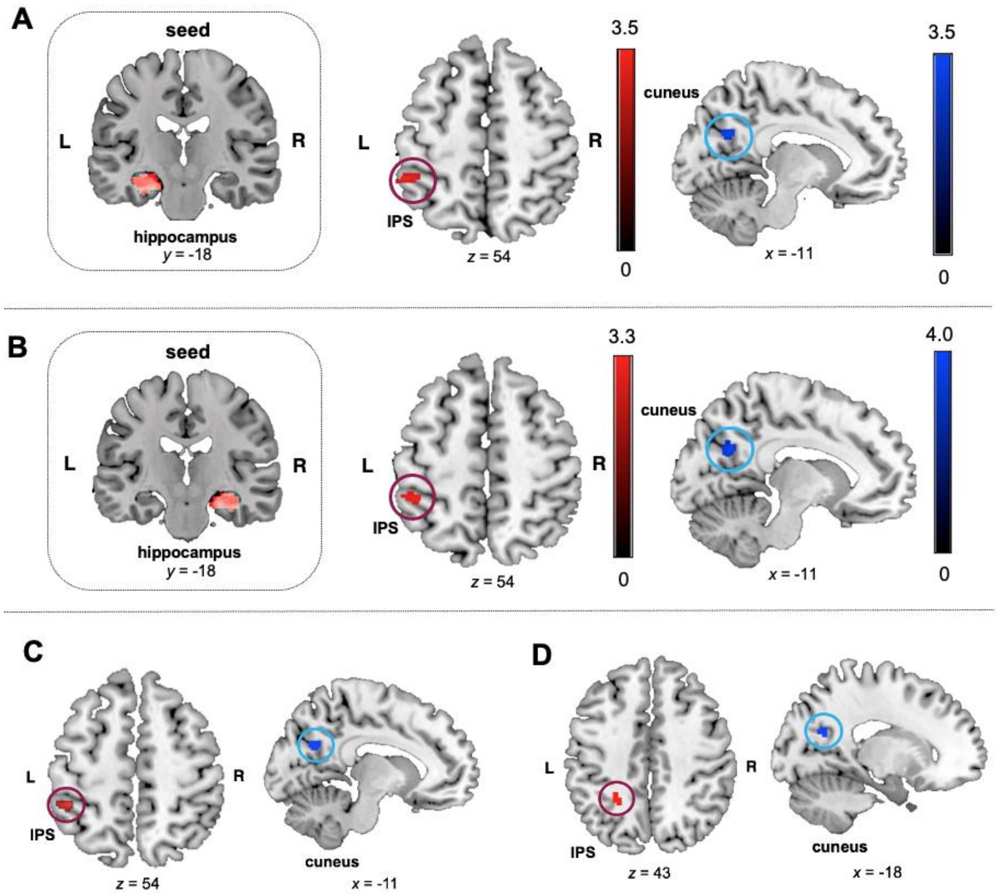

Figure 3. Hippocampal functional connectivity predicts number sense training gains. **A, B.** Functional connectivity before training of the left and right hippocampal regions of interest (ROIs; selected from Brainnetome (Fan et al., 2016) parcellations) with the left intraparietal sulcus (IPS) positively (red) predicted learning (efficiency gains in symbolic quantity discrimination) in the training group. Hippocampal connectivity with the left cuneus was negatively (blue) correlated with learning. **C**. Conjunction analysis of functional connectivity patterns revealed the left IPS and the left cuneus to be an overlapping target region across the left and right hippocampal seed-to-whole brain connectivity positively (red) and negatively (blue) associated with learning, respectively, in the training group. **D**. In typically developing (TD) children in the training group, hippocampal functional connectivity with the left IPS was positively associated with learning (red) and that with the left cuneus was negatively associated with learning (blue). In children with mathematical learning difficulties (MLD) in the training

group, conjunction analysis did not yield any overlapping target region across the left and right hippocampal seed-to-whole-brain functional connectivity associated with learning. L = left, R = right.

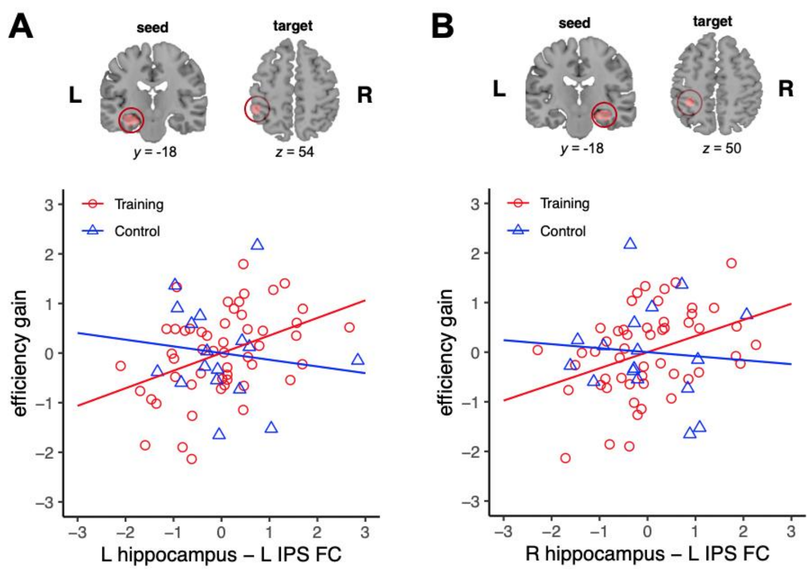

Figure 4. Hippocampal–parietal functional circuits predict number sense training gains. Scatter plots of the relationship between functional connectivity (FC) of the left (**A**) and right (**B**) hippocampus with the left intraparietal sulcus (IPS) and changes in efficiency in symbolic quantity discrimination (efficiency gain) in children who received training (Training) and no-contact control group (Control). Greater FC with the left IPS predicts efficiency gain in the Training group (ρs > .40, *p*s < .004), but not in the Control group (|ρ|s < .21, *p*s > .44). Target (IPS) regions of interest were identified using a 6-mm sphere centered on the peak voxel to estimate FC. L = left, R = right.

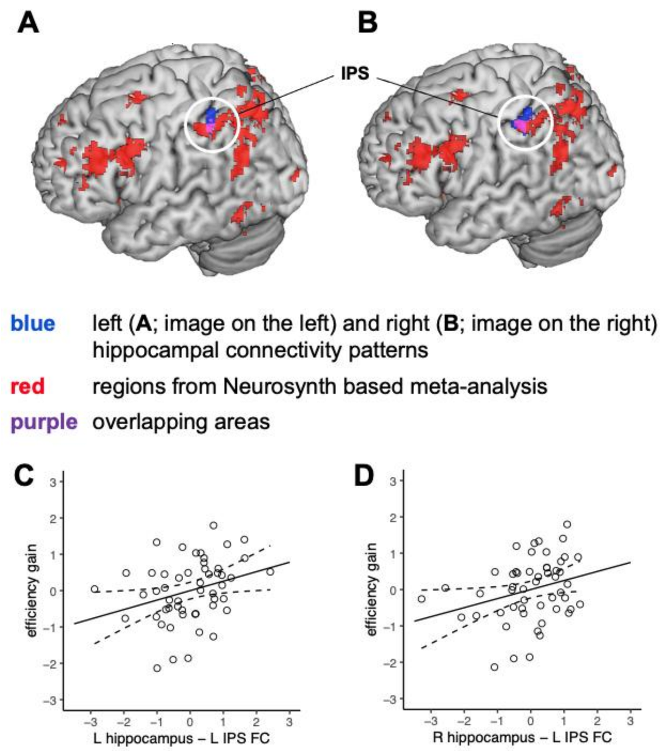

Figure 5. Hippocampal connectivity with a left intraparietal sulcus (IPS) region, identified using Neurosynth based meta-analysis, predicts number sense training gains. Left IPS regions (in white circle) overlap across (**A**) left and (**B**) right hippocampal connectivity patterns identified in the current study and the IPS region identified from Neurosynth based meta-analysis. (**C**-**D**). Scatter plots of the relationship between functional connectivity (FC) of the (**A**) left and (**B**) right hippocampal with the left IPS region identified from Neurosynth based meta-analysis. FC between the hippocampus and left IPS correlates with changes in efficiency in symbolic quantity discrimination (efficiency gain) following number sense training (ρs > .29, *p*s < .04). The IPS region of interest was identified using a 6-mm sphere centered on the peak voxel to estimate FC. L = left, R = right.

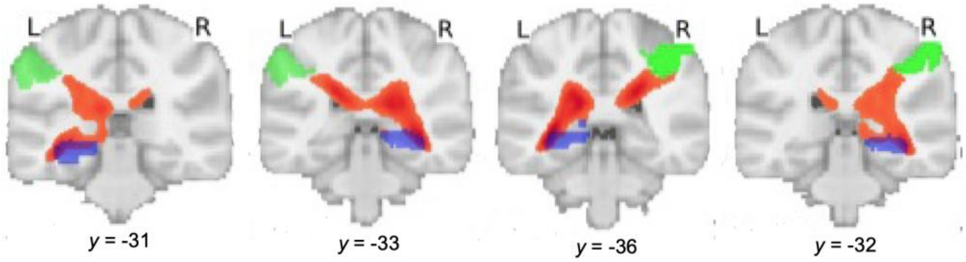

Figure 6. Hippocampal–parietal white matter tracts in children.

White matter tracts (depicted in orange) are identified between the hippocampus (purple) and intraparietal sulcus (IPS) (light green), averaged across all children using probabilistic tractography of High Angular Resolution Diffusion Imaging data. IPS regions of interest (ROIs) were selected from Brainnetome (Fan et al., 2016) parcellations that overlapped with target ROIs identified in functional connectivity analysis (left IPS) and contralateral regions (right IPS) to estimate hippocampal–parietal structural connectivity. L = left, R = right.

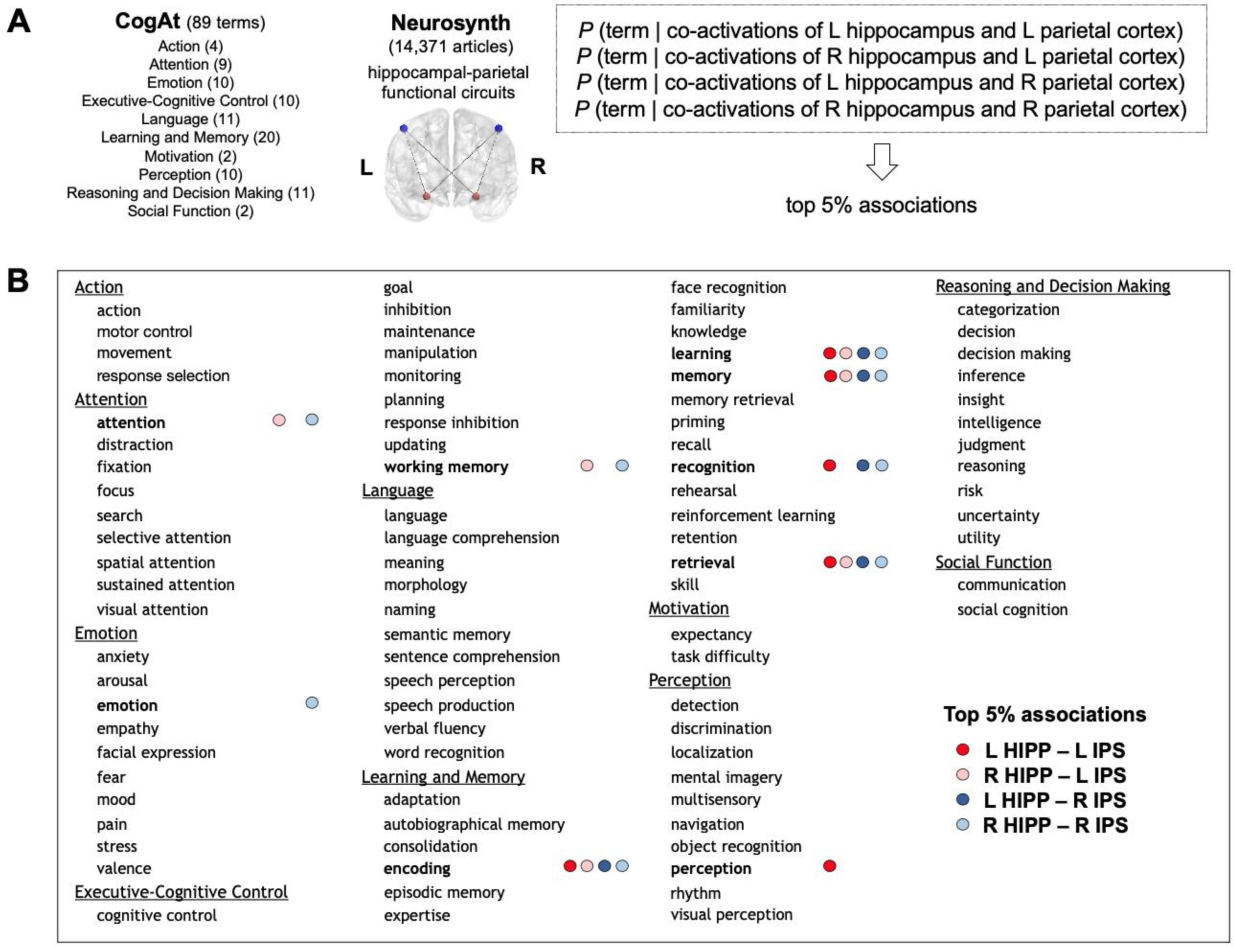

Figure 7. Reverse meta-analysis of 14,371 fMRI studies and cognitive functions reveals a significant association between hippocampal–parietal functional circuits and learning. **A.** A reverse meta-analysis was performed to map hippocampal–parietal functional circuits identified in the current study to cognitive functions (see *Methods* for details). **B.** Top 5% cognitive functions that are mentioned in published articles where co-activations of the left or right hippocampus (HIPP) and the left or right intraparietal cortex (IPS) are reported. L = left, R = right.

Table 1. Training activities in each session (3 sessions/week).

	Lessons	Games	Review
Week 1	• Counting • Video of a sock puppet counting	• Restaurant Game (Blair, 2013a, 2013b)	• Counting
Week 2	• Comparison • Math Circles	• Number Race (Wilson et al., 2006) • Math War (Iuculano et al., 2015) • Comparing Speed	• Comparison between non-symbolic quantities
Week 3	• Comparison • Math Circles	• Number Race (Wilson et al., 2006) • Math War (Iuculano et al., 2015) • Comparing Speed	• Comparison between non-symbolic and symbolic quantities
Week 4	• Comparison • Beat Your Score (Chang et al., 2019)	• Number Race (Wilson et al., 2006) • Math War (Iuculano et al., 2015) • Comparing Speed	• Comparison between symbolic quantities

Table 2. *Resting state fMRI data analysis sample*. Time 1 demographics and neuropsychological measures of 4-week number sense training (Training) and no-contact control (Control) groups.

	Training	Control	χ^2 test or two-sample *t*-test			
	M (SD)	*M (SD)*	${\chi^2}_1$ or *t*(67)	*p*	φ or Cohen's d	BF
Female to Male ratio	27 : 25	11 : 6	.41	.523	.08	.38
Age	8.21 (.61)	8.32 (.56)	.66	.513	.18	.33
WASI:						
Verbal IQ	109.25 (12.81)	107.35 (11.59)	-.54	.590	-.15	.32
Performance IQ	105.44 (14.73)	105.53 (11.41)	.02	.982	.01	.28
Full-Scale IQ	108.04 (12.78)	107.29 (10.02)	-.22	.828	-.06	.29
WJ-III:						
Math Fluency	97.40 (11.56)	101.41 (9.08)	1.30	.198	.36	.56
Calculation	104.90 (14.44)	106.88 (11.87)	.51	.611	.14	.31
Applied Problems	104.42 (13.15)	106.59 (10.88)	.61	.542	.17	.33
Letter-Word Identification	109.35 (9.40)	112.00 (8.84)	1.02	.309	.29	.43
Word Attack	106.73 (9.42)	107.41 (6.98)	.27	.785	.08	.29

Abbreviations: WASI = Wechsler Abbreviated Scale of Intelligence; WJ-III = Woodcock Johnson III.

Table 3. *Resting state fMRI data analysis sample.* Time 1 demographics neuropsychological measures of typically developing (TD) children and children with mathematical learning difficulties (MLD) in the training group (TD_Training and MLD_Training).

	TD_Training	MLD_Training	χ^2 test or two-sample *t*-test			
	M (SD)	*M (SD)*	${\chi^2}_1$ or $t(50)$	*p*	φ or Cohen's d	BF
Female to Male ratio	17 : 17	10 : 8	.01	.929	.01	.34
Age	8.12 (.57)	8.38 (.68)	1.47	.148	.43	.69
WASI:						
Verbal IQ	109.85 (12.23)	108.11 (14.14)	-.46	.646	-.13	.32
Performance IQ	107.24 (13.62)	102.06 (16.50)	-1.21	.231	-.35	.52
Full-Scale IQ	109.44 (11.70)	105.39 (14.59)	-1.09	.281	-.32	.47
WJ-III:						
Math Fluency	103.65 (9.09)	85.61 (3.97)	-8.00	<.001	-2.33	>100
Calculation	109.18 (14.46)	96.83 (10.70)	-3.18	.002	-.93	14.58
Applied Problems	109.06 (10.08)	95.67 (14.04)	-3.97	<.001	-1.16	105.31
Letter-Word Identification	111.21 (9.00)	105.83 (9.38)	-2.02	.049	-.59	1.47
Word Attack	108.88 (9.62)	102.67 (7.73)	-2.36	.022	-.69	2.65

Abbreviations: WASI = Wechsler Abbreviated Scale of Intelligence; WJ-III = Woodcock Johnson III.

Table 4. *Diffusion MRI data analysis sample*. Time 1 demographics and neuropsychological measures of 4-week number sense training (Training) and no-contact control (Control) groups.

	Training	Control	χ^2 test or two-sample *t*-test			
	M (SD)	*M (SD)*	χ^2_1 or $t(54)$	p	φ or Cohen's d	BF
Female to Male ratio	22 : 21	7 : 6	<.01	>.999	<.01	.28
Age	8.21 (.64)	8.39 (.56)	.92	.360	.29	.43
WASI:						
Verbal IQ	109.35 (12.85)	109.15 (12.12)	-.05	.962	-.02	.31
Performance IQ	103.95 (14.05)	106.23 (11.48)	.53	.597	.17	.35
Full-Scale IQ	107.30 (12.82)	108.54 (10.37)	.32	.752	.10	.32
WJ-III:						
Math Fluency	97.67 (11.66)	101.77 (8.76)	1.17	.248	.37	.53
Calculation	105.26 (13.41)	107.77 (12.09)	.61	.548	.19	.36
Applied Problems	103.88 (12.08)	109.85 (10.16)	1.61	.113	.51	.86
Letter-Word Identification	108.93 (9.13)	111.00 (8.42)	.73	.469	.23	.38
Word Attack	106.58 (9.65)	106.38 (6.32)	-.07	.945	-.02	.31

Abbreviations: WASI = Wechsler Abbreviated Scale of Intelligence; WJ-III = Woodcock Johnson III.

Table 5. *Diffusion MRI data analysis sample*. Time 1 demographics neuropsychological measures of typically developing (TD) children and children with mathematical learning difficulties (MLD) in the training group (TD_Training and MLD_Training).

	TD_Training	MLD_Training	χ^2 test or two-sample *t*-test			
	M (SD)	*M (SD)*	${\chi^2}_1$ or $t(41)$	*p*	φ or Cohen's d	BF
Female to Male ratio	15 : 13	7 : 8	.01	.911	.02	.38
Age	8.15 (0.61)	8.33 (0.70)	.91	.370	.29	.43
WASI:						
Verbal IQ	109.00 (12.11)	110.00 (14.55)	.24	.811	.08	.32
Performance IQ	105.86 (13.69)	100.40 (14.51)	-1.22	.229	-.39	.56
Full-Scale IQ	108.25 (11.98)	105.53 (14.54)	-.66	.514	-.21	.37
WJ-III:						
Math Fluency	103.86 (9.56)	86.13 (3.50)	-6.90	<.001	-2.21	>100
Calculation	108.71 (14.11)	98.80 (9.30)	-2.44	.019	-.78	3.05
Applied Problems	108.46 (10.05)	95.33 (11.09)	-3.94	<.001	-1.26	81.81
Letter-Word Identification	110.82 (8.82)	105.40 (8.93)	-1.91	.063	-.61	1.29
Word Attack	108.86 (10.38)	102.33 (6.47)	-2.21	.033	-.71	2.03

Abbreviations: WASI = Wechsler Abbreviated Scale of Intelligence; WJ-III = Woodcock Johnson III.

Table 6. Brain regions showing positive and negative relation between functional connectivity with the left hippocampus and symbolic quantity discrimination efficiency gain in response to 4-week number sense training.

Region	Number of voxels	Peak intensity	MNI coordinates: x,y,z (mm)
Positive relation			
Training			
L IPS/SMG/SPL	83	3.49	-48 -42 54
Training: TD children			
L IPS/SPL/SMG	69	4.59	-24 -48 42
L MFG/PCG/SFG	68	4.18	-22 12 50
Training: children with MLD			
(none)			
Negative relation			
Training			
L CUN/PCUN	112	3.48	-10 -64 26
Training: TD children			
L SOG/CUN/PCUN	243	4.46	-22 -62 24
Training: children with MLD			
(none)			

Notes: Anatomical locations of brain regions were identified by Automated Anatomical Labeling (AAL) (Tzourio-Mazoyer et al., 2002), Harvard-Oxford (Desikan et al., 2006), and Juelich histological (Eickhoff et al., 2005) atlases. *Abbreviations*: CUN = cuneus; IPS = intraparietal sulcus; MFG = middle frontal gyrus; PCG = precentral gyrus; PCUN = precuneus; SFG = superior frontal gyrus; SMG = supramarginal gyrus; SOG = superior occipital gyrus; SPL =

superior parietal lobule; TD = typically developing; MLD = mathematical learning difficulties; L = left; R = right.

Table 7. Brain regions showing positive and negative relation between functional connectivity with the right hippocampus and symbolic quantity discrimination efficiency gain in response to 4-week number sense training.

Region	Number of voxels	Peak intensity	MNI coordinates: x,y,z (mm)
Positive relation			
Training			
L IPS/SPL/SMG	156	3.33	-34 -40 50
Training: TD children			
L IPS/SPL/SMG	75	4.20	-24 -48 44
Training: children with MLD			
(none)			
Negative relation			
Training			
L CUN/PCUN	140	3.92	-12 -62 28
Training: TD children			
L PCUN/PCC	86	4.69	-4 -50 10
L SOG/CUN/PCUN	138	3.95	-22 -62 24
R SMG/AG/IPS	219	3.71	38 -52 28
L AG/LOC	147	3.44	-40 -76 44
Training: children with MLD			
R CUN/LOC/OP	314	7.56	14 -90 30

Notes: Anatomical locations of brain regions were identified by Automated Anatomical Labeling (AAL) (Tzourio-Mazoyer et al., 2002), Harvard-Oxford (Desikan et al., 2006), and Juelich histological (Eickhoff et al., 2005) atlases. *Abbreviations*: AG = angular gyrus; CUN = cuneus; IPS = intraparietal sulcus; LOC = lateral occipital cortex; OP = occipital pole; PCUN =

precuneus; SMG = supramarginal gyrus; SOG = superior occipital gyrus; SPL = superior parietal lobule; TD = typically developing; MLD = mathematical learning difficulties; L = left; R = right.

Table 8. Correlations between functional and structural connectivity and symbolic quantity discrimination efficiency gain in 4-week number sense training (Training) and no-contact control (Control) groups.

	Training (*N* = 52[†]; 43[‡])			**Control** (*N* = 17[†]; 13[‡])		
	ρ	*p*	BF	ρ	*p*	BF
Functional connectivity						
L HIPP - L IPS	.42**	.002	**18.10**	-.20	.445	.57
R HIPP - L IPS	.41**	.003	**9.93**	-.14	.589	.53
L HIPP - R IPS	-.10	.500	.33	-.14	.586	.54
R HIPP - R IPS	.15	.289	.34	-.10	.708	.55
Structural connectivity						
L HIPP - L IPS	-.04	.821	.39	.16	.603	.57
R HIPP - L IPS	-.20	.190	.61	-.08	.803	.75
L HIPP - R IPS	-.12	.448	.54	.08	.803	.59
R HIPP - R IPS	-.03	.865	.38	-.25	.415	1.10

Abbreviations: HIPP = hippocampus; IPS = intraparietal sulcus; BF = Bayes factor; L = left; R = right. *Notes*: [†]number of participants included in functional connectivity analysis. [‡]number of participants included in structural connectivity analysis. ** $p < .01$. Bolded BF values (>3) provide evidence for H_1. BF values between .33 and 3 provide absence of evidence (i.e., insufficient evidence for either H_1 or H_o) (Keysers et al., 2020).

Table 9. Correlations between functional and structural connectivity and symbolic quantity discrimination efficiency gain in typically developing (TD) children and children with mathematical learning difficulties (MLD) in the training group (TD_Training and MLD_Training).

	TD_Training (*N* = 34[†]; 28[‡])			**MLD_Training** (*N* = 18[†]; 15[‡])		
	ρ	*p*	BF	ρ	*p*	BF
Functional connectivity						
L HIPP - L IPS	.43*	.012	**3.98**	.52*	.026	2.20
R HIPP - L IPS	.38*	.025	**4.58**	.52*	.027	.93
L HIPP - R IPS	-.16	.358	.44	<.01	.997	.50
R HIPP - R IPS	.03	.855	.39	.32	.197	.54
Structural connectivity						
L HIPP - L IPS	-.07	.734	.42	-.11	.694	.59
R HIPP - L IPS	-.23	.238	1.33	-.25	.362	.54
L HIPP - R IPS	-.05	.806	.43	-.38	.164	.91
R HIPP - R IPS	.03	.897	.41	-.24	.398	.57

Abbreviations: HIPP = hippocampus; IPS = intraparietal sulcus; BF = Bayes factor; L = left; R = right. *Notes*: [†]number of participants included in functional connectivity analysis. [‡]number of participants included in structural connectivity analysis. * $p < .05$. Bolded BF values (>3) provide evidence for H_1. BF values between .33 and 3 provide absence of evidence (i.e., insufficient evidence for either H_1 or H_o) (Keysers et al., 2020).

Table 10. Model comparison between multiple regression analysis including functional and/or structural connectivity measures as predictors of symbolic quantity discrimination efficiency gain in the training group (N = 43).

Model fit	Adjusted R^2	F	df	p
Model 1: Baseline (1 + Control)	.25	14.73***	1, 41	<.001
Model 2: (Baseline + FC)	.44	7.50***	5, 37	<.001
Model 3: (Baseline + SC)	.22	3.33*	5, 37	.014
Model 4: (Baseline + FC + SC)	.54	6.55***	9, 33	<.001

Model comparison	ΔR^2	F	df	p	BF
Model 2 (Model 1 + FC) vs. Model 1 (Baseline)	.24	4.45**	4	.005	**7.33**
Model 3 (Model 1 + SC) vs. Model 1	.05	.62	4	.65	.05
Model 4 (Model 2 + SC) vs. Model 2	.14	3.17*	4	.026	2.26
Model 4 (Model 3 + FC) vs. Model 3	.33	7.61***	4	<.001	**>100**

Notes: Control = Time 1 symbolic quantity discrimination task efficiency; FC = estimates of ipsilateral and contralateral functional connectivity between left and right hippocampus and left and right intraparietal sulcus; SC = estimates of ipsilateral and contralateral structural connectivity between hippocampus and intraparietal sulcus. * p < .05; ** p < .01; *** p <.001. Bolded Bayes Factor (BF) (>3) provides evidence for H_1. BF values between .33 and 3 provide absence of evidence (i.e., insufficient evidence for either H_1 or H_o). BF values below .33 provide evidence of absence (evidence for H_o) (Keysers et al., 2020).