



**HAL**  
open science

## 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

► **To cite this version:**

Hyesang Chang, Lang Chen, Yuan Zhang, Ye Xie, Carlo de Los Angeles, et al.. Foundational number sense training gains are predicted by hippocampal–parietal circuits. *Journal of Neuroscience*, 2022, 42 (19), pp.JN-RM-1005-21. 10.1523/JNEUROSCI.1005-21.2022 . hal-03670987

**HAL Id: hal-03670987**

**<https://inria.hal.science/hal-03670987v1>**

Submitted on 18 May 2022

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

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

**Abbreviated title:** Learning related hippocampal–parietal circuits

*Hyesang Chang<sup>1\*</sup>, Lang Chen<sup>1,2</sup>, Yuan Zhang<sup>1</sup>, Ye Xie<sup>1,3,4</sup>, Carlo de Los Angeles<sup>1</sup>, Emma Adair<sup>1</sup>,  
Gaston Zanitt<sup>5</sup>, Demian Wassermann<sup>5</sup>, Miriam Rosenberg-Lee<sup>1,6</sup>, Vinod Menon<sup>1,7,8\*</sup>*

<sup>1</sup>Department of Psychiatry & Behavioral Sciences, Stanford University School of Medicine,  
Stanford, CA 94305, US

<sup>2</sup>Department of Psychology, Santa Clara University, Santa Clara, CA 95053, US

<sup>3</sup>Department of Physics, Zhejiang University, Hangzhou 310027, China

<sup>4</sup>Department of Psychology, Sun Yat-Sen University, Guangzhou 510006, China

<sup>5</sup>Parietal, Inria Saclay Île-de-France, CEA Université Paris Sud 1 Rue Honoré d'Estienne  
d'Orves 91120, Palaiseau, France

<sup>6</sup>Department of Psychology, Rutgers University, Newark, NJ 07102, US

<sup>7</sup>Department of Neurology & Neurological Sciences,

<sup>8</sup>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](mailto:changh@stanford.edu); [menon@stanford.edu](mailto: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.

1 **Abstract**

2 The development of mathematical skills in early childhood relies on number sense, the  
3 foundational ability to discriminate between quantities. Number sense in early childhood is  
4 predictive of academic and professional success, and deficits in number sense are thought to  
5 underlie lifelong impairments in mathematical abilities. Despite its importance, the brain circuit  
6 mechanisms that support number sense learning remain poorly understood. Here, we designed  
7 a theoretically motivated training program to determine brain circuit mechanisms underlying  
8 foundational number sense learning in female and male elementary school-aged children (ages  
9 7-10). Our four-week integrative number sense training program gradually strengthened the  
10 understanding of the relations between symbolic (Arabic numerals) and non-symbolic (sets of  
11 items) representations of quantity. We found that our number sense training program improved  
12 symbolic quantity discrimination ability in children across a wide a range of math abilities  
13 including those with learning difficulties. Crucially, the strength of pre-training functional  
14 connectivity between the hippocampus and intraparietal sulcus, brain regions implicated in  
15 associative learning and quantity discrimination, respectively, predicted individual differences in  
16 number sense learning across typically developing children and children with learning difficulties.  
17 Reverse meta-analysis of inter-regional co-activations across 14,371 fMRI studies and 89  
18 cognitive functions confirmed a reliable role for hippocampal–intraparietal-sulcus circuits in  
19 learning. Our study identifies a canonical hippocampal–parietal circuit for learning which plays a  
20 foundational role in children’s cognitive skill acquisition. Findings provide important insights into  
21 neurobiological circuit markers of individual differences in children’s learning and delineate a  
22 robust target for effective cognitive interventions.

23 **Significance Statement**

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

### 33 **Introduction**

34 Number sense, the ability to discriminate quantities, is predictive of academic achievement and  
35 professional success (Jordan, Kaplan, Ramineni, & Locuniak, 2009; National Mathematics  
36 Advisory Panel, 2008). In particular, weaknesses in mapping symbolic numbers (e.g., the  
37 symbol “2”) to their magnitude representations (e.g., two objects) in early childhood are  
38 associated with difficulties in subsequent mathematical skill acquisition (De Smedt & Gilmore,  
39 2011; Rousselle & Noël, 2007). The knowledge about the neural mechanisms that support  
40 number sense acquisition in children can provide important insights into neurobiological markers  
41 of individual differences in learning and inform more effective interventions. Here we develop a  
42 theoretically motivated training protocol to address a critical gap in our understanding of the  
43 acquisition of foundational skills and the brain circuits that predict learning in early elementary  
44 school children.

45  
46 Recent studies have begun to uncover a crucial role for the hippocampal memory system in  
47 mathematical skill acquisition (Menon & Chang, 2021; Supekar, Chang, Mistry, Iuculano, &  
48 Menon, 2021). Most previous studies of mathematical cognition have focused on cortical  
49 regions, most notably the posterior parietal cortex; consequently, the role of medial temporal  
50 lobe regions in mathematical cognition and learning has received less attention because of this  
51 cortico-centric focus. Notably, in a study on arithmetic fact learning (e.g.,  $3 + 5 = 8$ ),  
52 hippocampal functional connectivity with multiple cortical regions predicted individual differences  
53 in performance gains in children (Supekar et al., 2013). Furthermore, the strength of this  
54 association with learning was stronger than the connectivity of the intraparietal sulcus (IPS), a  
55 brain region consistently implicated in representation of numerical quantities across symbolic  
56 and non-symbolic formats (Butterworth & Walsh, 2011; Piazza & Eger, 2016). These findings  
57 provide support for theoretical models which posit that hippocampal circuitry is critical during  
58 early stages of math skill acquisition (Menon, 2016; Menon & Chang, 2021). Here we test the

59 hypothesis that brain circuits linking the hippocampus, a brain region crucial for binding new  
60 information (Eichenbaum, 2004; Olsen, Moses, Riggs, & Ryan, 2012; Zeithamova & Bowman,  
61 2020), and parietal cortical areas important for representation of numerical quantity supports  
62 foundational number sense learning.

63

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

85 across 14,371 published fMRI studies in relation to 89 cognitive functions to determine whether  
86 hippocampal–parietal functional circuits identified in the present study constitute a canonical  
87 circuit for learning.

88

## 89 **Materials and Methods**

### 90 ***Participants***

91 A total of 96 children in 2<sup>nd</sup> and 3<sup>rd</sup> grades (age:  $M = 8.19$ ,  $SD = 0.63$ , 54 females) recruited with  
92 flyers sent to schools and posted at libraries and community centers in the San Francisco Bay  
93 Area participated in the study. All participants were right-handed and did not report any current  
94 neurological or psychiatric illness. Among them, 66 children (age:  $M = 8.15$ ,  $SD = 0.65$ , 35  
95 females) participated in the training program and 30 children (age:  $M = 8.27$ ,  $SD = 0.61$ , 19  
96 females) served as no-contact controls. MLD in children was identified using *normed-based*  
97 *cutoff criteria* applied to math fluency, similar to previously published studies (Iuculano et al.,  
98 2015; Jolles et al., 2016; Rosenberg-Lee et al., 2015). Children who scored at or below 90 (25<sup>th</sup>  
99 percentile or below) on the Math Fluency subtest of the Woodcock Johnson Third Edition (WJ-III)  
100 (Woodcock, McGrew, & Mather, 2001) were included in the MLD group ( $M = 85.61$ ,  $SD = 3.97$ ),  
101 and children who scored above 90 were included in the TD group ( $M = 103.65$ ,  $SD = 9.09$ ). No  
102 participant was excluded due to MLD or TD status.

103

104 All study protocols were approved by the Stanford University School of Medicine Institutional  
105 Review board and informed consent was obtained from the parents of the children. Children  
106 received \$50 for completing each MRI scanning session, \$50 for completing neuropsychological  
107 assessment battery, and \$50 for participating in the training program. Twenty-seven children  
108 were excluded from resting-state fMRI analysis due to no structural MRI data acquired ( $n = 3$ ),  
109 poor structural MRI image quality ( $n = 6$ ), missing behavioral data from fMRI task ( $n = 2$ ), or  
110 inadequate whole brain coverage or excessive head movement ( $n = 16$ ; see fMRI

111 preprocessing in *Methods*) in the scanner. For diffusion MRI analysis, 13 children were  
112 additionally excluded due to incomplete diffusion MRI data acquisition ( $n = 1$ ), poor diffusion  
113 MRI data quality ( $n = 10$ ), or identification as extreme outliers in structural connectivity  
114 measures ( $n = 2$ ). The final resting-state fMRI analysis sample included 52 children (18 children  
115 with MLD, 34 TD children) and 17 children in the training and control group, respectively; the  
116 diffusion MRI analysis sample included 43 children (15 children with MLD, 28 TD children) and  
117 13 children in the training and control group, respectively.

118

### 119 ***Experimental design and statistical analyses***

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

127

128 Children in the training group completed a 4-week number sense training program (3 days/week,  
129 for approx. 60 minutes/day), which focused on strengthening of children’s understanding of the  
130 relations between symbolic and non-symbolic representations of quantity ranging from 1 to 9.  
131 The first week of training began with a review of counting principles, followed by practice in  
132 enumeration and comparisons between non-symbolic quantities (sets of items or dot arrays) in  
133 the second week, comparisons between non-symbolic and symbolic (Arabic numbers)  
134 quantities in the third week, and finally, comparisons between symbolic quantities in the fourth  
135 week. Training with non-symbolic numbers included in weeks 2 and 3 was used to scaffold  
136 children’s learning of symbolic numbers through mapping non-symbolic quantities to verbal

137 (number words) or visual (Arabic numerals) symbolic numbers. This was followed by training  
138 with symbolic numbers, without the use of non-symbolic numbers, in the final week. Response  
139 to training was examined using symbolic quantity discrimination task acquired before and after  
140 training.

141

142 For statistical analyses of behavioral data, two-sample  $t$ -tests, Chi squared tests, and  
143 multivariate analysis of variance (MANOVA) were performed for comparisons between groups  
144 of age, gender, or neuropsychological assessments. A repeated measures ANOVA with time as  
145 a within-subject factor and group as a between-subject factor was conducted to assess the  
146 effects of training on symbolic quantity discrimination task performance. Follow-up paired  $t$ -tests  
147 examined changes in task performance (learning gains) in each group. In addition, two sample  
148  $t$ -tests assessed differences in learning gains and pre- and post-training differences in task  
149 performance between groups. Spearman correlations were used for analysis on relations  
150 between behavioral measures and brain-behavior relations to minimize influence of potential  
151 outliers. In addition to frequentist statistics (e.g.,  $p$ -values), Bayes factor (BF) was used to  
152 assess presence or absence of evidence for  $H_1$  or  $H_0$  (Keysers, Gazzola, & Wagenmakers,  
153 2020). BF values greater than 3 provide evidence for  $H_1$ . BF values between .33 and 3 provide  
154 absence of evidence. BF values below .33 provide evidence of absence (evidence for  $H_0$ ).

155

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

161

162 ***Training sessions***

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

168  
169 **Lessons.** In tutoring sessions in week 1, the tutor provided a lesson on counting by reviewing  
170 counting principles with examples of correct and incorrect counting of erasers; after the lesson,  
171 children viewed a video of a sock puppet counting and were asked to determine whether or not  
172 the sock puppet counted correctly. In tutoring sessions from weeks 2 to 4, children were asked  
173 to count out loud from 1 to 9 and was reminded that each number they counted up was bigger  
174 than the number before it. Then, in weeks 2 and 3, children completed *Math Circles*, where they  
175 enumerated the number (of erasers) in each of two *Math Circles* on the table and determined  
176 which number is bigger than the other. In week 2 (lesson on non-symbolic quantities), the tutor  
177 put two sets of erasers of different quantities in the two *Math Circles* (one quantity in each Math  
178 Circle). In week 3 (lesson on non-symbolic and symbolic quantities), the tutor put a card  
179 showing a number (Arabic numeral) and a set of erasers of different quantities from the number  
180 on the card in the two *Math Circles*. In week 4 (lesson on symbolic quantities), the tutor  
181 administered a number-ordering version of *Beat Your Score* (Chang, Rosenberg-Lee, Qin, &  
182 Menon, 2019), where the child was asked to order 4 decks of cards with quantities in non-  
183 symbolic (array of dots), mixed (symbolic and non-symbolic), or symbolic (Arabic numeral)  
184 format more quickly each time they ordered the cards. For each deck of cards, children  
185 completed the ordering of cards 3 times (after the tutor shuffled the cards), with the goal of  
186 beating their previous time taken to order the cards on the 3<sup>rd</sup> time for at least 3 decks of cards.  
187

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

210

211 **Review worksheets.** The review worksheets consisted of counting the number of animals  
212 (week 1) or identifying the larger quantity between two non-symbolic (week 2), symbolic and  
213 non-symbolic (week 3), or symbolic (week 4) quantities. Children circled, matched, or wrote the

214 number of animals (week 1) or circled the larger quantity (weeks 2-4) on the worksheet.  
215 Accuracy was emphasized in the worksheets from weeks 1 to 3 (worksheet included 42 trials in  
216 week 1, 24 trials in week 2, and 48 trials in week 3). In week 4, children were given one minute  
217 to complete 24 symbolic number comparison trials the worksheet.

218

### 219 ***Symbolic quantity discrimination task***

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

234

### 235 ***MRI data acquisition***

236 Images were acquired on a 3T GE Signa scanner (General Electric, Milwaukee, WI) using a  
237 custom-built head coil at the Stanford University Lucas Center. Head movement was minimized  
238 during the scan by cushions placed around the participant's head. A total of 31 axial slices (4.0  
239 mm thickness, 0.5 mm skip) parallel to the anterior commissure (AC)-posterior commissure (PC)

240 line and covering the whole brain were imaged using a T2\* weighted gradient echo spiral in-out  
241 pulse sequence (Glover & Lai, 1998) with the following parameters: TR = 2 sec, TE = 30 msec,  
242 flip angle = 80°, 1 interleave. The field of view was 22 cm, and the matrix size was 64 x 64,  
243 providing an in-plane spatial resolution of 3.4375 mm. The total length of the run was 6 minutes  
244 and 10 seconds. To reduce blurring and signal loss from field inhomogeneity, an automated  
245 high-order shimming method based on spiral acquisitions was used before acquiring functional  
246 MRI scans (Kim, Adalsteinsson, Glover, & Spielman, 2002). High-resolution T1-weighted 3D  
247 MRI sequences were acquired to facilitate anatomical co-registration of fMRI maps, with the  
248 following parameters: I = 400ms, TR = 5.9ms; TE = minimum; flip angle = 11°; field of view =  
249 240mm; matrix size = 256 x 192; 170 axial slices (1.0-mm thickness).

250  
251 A state-of-the-art diffusion-weighted single-shot spin-echo, echo planar imaging HARDI pulse  
252 sequence was used for more precise examination of white-matter fibers, including crossing  
253 fibers (Tuch et al., 2002), with the following parameters: TR = 5.3s; TE = minimum; flip angle =  
254 90°; field of view = 260mm; matrix size = 128 x 128; 50 axial slices (2.9-mm thickness, no  
255 spacing). The high b value (2500s/mm<sup>2</sup>) was obtained by applying gradients along 150 different  
256 diffusion directions.

257

### 258 ***fMRI preprocessing***

259 Resting-state functional MRI data were analyzed using SPM12  
260 (<http://www.fil.ion.ucl.ac.uk/spm/>). The first 5 volumes were not analyzed to allow for T1  
261 equilibration. A linear shim correction was applied separately for each slice during  
262 reconstruction (Glover & Lai, 1998). Images were realigned to the first scan to correct for motion  
263 and slice acquisition timing, co-registered to each individual's structural T1 images, spatially  
264 transformed to standard stereotaxic space (based on the Montreal Neurologic Institute  
265 coordinate system), resampled every 2 mm using sinc interpolation, and smoothed with a 6mm

266 full-width half maximum Gaussian kernel to decrease spatial noise prior to statistical analysis. A  
267 bandpass filter (.008-.1 Hz) was applied to the smoothed data to remove high frequency  
268 artifacts. Translational movement in millimeters (x,y,z), and rotational motion in degrees (pitch,  
269 roll, yaw) were calculated based on the SPM12 parameters for motion correction of the  
270 functional images of each subject. We excluded participants with movement larger than 10mm  
271 in any of the x,y,z directions. Mean scan-to-scan displacement of movement did not exceed  
272 0.5mm for all participants.

273

#### 274 ***Intrinsic functional connectivity analysis***

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

286

287 Pre-training performance on symbolic quantity discrimination was regressed out to control for  
288 regression to the mean, a known phenomenon in intervention studies (Barnett, van der Pols, &  
289 Dobson, 2005), and to obtain more precise estimates of intervention effects (Pocock, Assmann,  
290 Enos, & Kasten, 2002; Thompson, Lingsma, Whiteley, Murray, & Steyerberg, 2015). This  
291 analysis approach allowed us to minimize the potential influence of pre-training performance on

292 learning. In the current study, the association between change in performance and pre-training  
293 performance on symbolic quantity discrimination was significant in the training ( $\rho = -.50$ ,  $p$   
294  $< .001$ ) but not in the control ( $\rho = -.18$ ,  $p = .49$ ) group, which indicates that the degree to which  
295 pre-training performance influences changes in performance varied between groups.

296

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

311

312 Follow-up ROI-based analyses were performed to visualize the results, ensure that the results  
313 were not driven by outliers, and confirm differences in correlation between brain and behavioral  
314 measures across groups. ROIs were defined as 6-mm spheres centered at peak of the left  
315 intraparietal sulcus (IPS) identified from hippocampal connectivity patterns in the training group  
316 and its contralateral region for the right IPS. Similar to the whole brain analysis (see above),

317 pre-training symbolic quantity discrimination task efficiency was regressed out from symbolic  
318 quantity discrimination task efficiency gains and connectivity estimates in each group.

319

### 320 ***Structural connectivity analysis***

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

343 strength between each pair of ROIs per subject. ROI-based analyses were then performed to  
344 assess the relation between hippocampal–parietal structural connectivity and learning, using  
345 similar procedures described in intrinsic functional connectivity analysis.

346

### 347 ***Cross-validation analysis***

348 A machine-learning approach with balanced fourfold cross-validation combined with linear  
349 regression (<https://github.com/poldrack/regressioncv>) (Cohen et al., 2010; Supekar et al., 2013)  
350 was conducted to investigate the robustness of brain-based predictors of individual differences  
351 in number sense learning gains. Learning gain as a dependent variable and the brain-based  
352 predictor (connectivity) as an independent variable were treated as inputs to a linear regression  
353 algorithm.  $r(\text{predicted, observed})$ , a measure of how well the independent variable predicts the  
354 dependent variable, was first estimated using a balanced fourfold cross-validation procedure.  
355 Participants were assigned to one of four folds. A linear regression model was built using three  
356 folds leaving out the fourth, and this model was then used to predict the data in the left-out fold.  
357 This procedure was repeated four times to compute a final  $r(\text{predicted, observed})$  representing  
358 the correlation between the data predicted by the regression model and the observed data.  
359 Finally, the statistical significance of the model was assessed using a non-parametric testing  
360 approach. The empirical null distribution of  $r(\text{predicted, observed})$  was estimated by generating  
361 1000 surrogate datasets under the null hypothesis that there was no association between  
362 changes in numerical skills and brain-based predictor.

363

### 364 ***Reverse meta-analysis***

365 To examine the role of the hippocampal–parietal functional circuits, we conducted a novel  
366 reverse meta-analysis of inter-regional coactivation of the hippocampus and the parietal cortex,  
367 using regions identified in the present study, reported across 14,371 published fMRI studies up  
368 to July 2018 from NeuroSynth (Yarkoni, Poldrack, Nichols, Van Essen, & Wager, 2011)

369 database and 89 cognitive atlas terms (CogAt) (Poldrack et al., 2011) (**Figure 7A**). The  
370 hippocampal ROIs were from Brainnetome (Fan et al., 2016) parcellations (rostral and caudal  
371 subdivisions combined) and the parietal ROIs were from Brainnetome (Fan et al., 2016)  
372 parcellations that overlapped with target ROIs identified from intrinsic functional connectivity  
373 analysis (superior parietal lobule [lateral area 5] and inferior parietal lobule [rostradorsal area 40]  
374 subdivisions combined; left IPS) and the contralateral regions (right IPS).

375

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

379

380  $P(\text{term 'learning' is mentioned} \mid \text{activations are reported in the left hippocampus and in the left}$   
381  $\text{parietal cortex})$

382

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

390

391 To estimate reverse meta-analysis probabilities, we followed the following steps. First, we  
392 encoded the probability of a term being present in a study by thresholding the term frequency –  
393 inverse document frequency (TF – IDF) value of the term being present at  $10^{-3}$ , in agreement  
394 with NeuroSynth’s implementation (Yarkoni et al., 2011). Second, we considered the probability

395 of a region being reported in a given study as directly proportional to the number of activations  
396 within the regions being present in the study, for which we resampled the activation foci to  $4\text{mm}^3$   
397 voxels in MNI152 space. Third, terms were filtered using the CogAt (Poldrack et al., 2011)  
398 ontology to ensure that only those relating to cognitive processes (89 terms listed in **Figure 7B**)  
399 were taken into account. To assess the stability of our estimations, we computed the confidence  
400 interval of our reverse meta-analysis probability estimations, we split the 14,371 studies in 20  
401 equal folds, maximizing the measurements for estimation. Finally, top 5% probable terms were  
402 considered to be sufficient evidence for associations with analyzed circuits.

403  
404 Our analysis resulted in the selection of 25 out of 356 associations (4 circuits, 89 terms), which  
405 represented above 95 percentile of probable term mentions for studies where hippocampal–  
406 parietal circuits were reported. For these top 5% terms, the maximum probability was estimated,  
407 across all splits, at  $0.34 \pm 0.011$  for *memory* being mentioned in studies where left hippocampus  
408 and right IPS activations are simultaneously reported, and the minimum was estimated at  
409  $0.10 \pm 0.005$  for mentioning *recognition* in studies where left hippocampus and left IPS  
410 activations are simultaneously reported. Hence, the standard deviation for all top 5%  
411 probabilities, is an order of magnitude smaller than the estimated probability, pointing out to a  
412 high confidence in our estimation.

413

## 414 **Results**

### 415 ***Comparison of neuropsychological measures between groups***

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

421

422 Children in the training and control groups did not significantly differ in age ( $t(67) = .66, p = .51,$   
423 Cohen's  $d = .18, BF = .33$ ) and gender ( $\chi^2_1 = .41, p = .52, \varphi = .08, BF = .38$ ; **Table 2**). A  
424 MANOVA between training and control groups on multiple neuropsychological assessments,  
425 including Wechsler Abbreviated Scale of Intelligence (WASI) (Wechsler, 1999) (Full-Scale,  
426 Verbal, and Performance IQ) and WJ-III (Math Fluency, Calculation, Applied Problems, Letter-  
427 Word Identification, and Word Attack) subtests ( $F(8,60) = .56, p = .807$ ) showed no significant  
428 difference between groups. Two-sample  $t$ -tests confirmed that children in the training and  
429 control groups did not significantly differ in WASI ( $|t|s < .55, ps > .58, |Cohen's d| < .16, BFs$   
430  $< .33$ ) and WJ-III subtests ( $ts < 1.31, ps > .19, Cohen's d < .37, BFs < .57$ ; **Table 2**).

431

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

442

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

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

454

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

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

467

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

471 = .21,  $p = .65$ ,  $\eta_p^2 < .01$ ). Training significantly improved symbolic quantity discrimination task  
472 efficiency in the training group ( $t(51) = 6.28$ ,  $p < .001$ , Cohen's  $d = .87$ ,  $BF > 100$ ), but not in the  
473 control group ( $t(16) = .17$ ,  $p = .86$ , Cohen's  $d = .04$ ,  $BF = .25$ , **Figure 2B**). Both TD children ( $t(33)$   
474 = 5.47,  $p < .001$ , Cohen's  $d = .94$ ,  $BF > 100$ ) and children with MLD ( $t(17) = 3.16$ ,  $p = .006$ ,  
475 Cohen's  $d = .74$ ,  $BF = 8.43$ ) improved with large individual differences in both groups  
476 (coefficient of variation: TD children: 1.06; children with MLD: 1.34). In addition, learning gains –  
477 changes (post-training – pre-training) in symbolic quantity discrimination task efficiency – were  
478 not significantly different between the two training groups (two-sample  $t$ -test,  $t(50) = -.26$ ,  $p = .80$ ,  
479 Cohen's  $d = -.08$ ,  $BF = .30$ ). These results demonstrate that four weeks of number sense  
480 training improved symbolic quantity discrimination ability in both TD children and children with  
481 MLD.

482  
483 Surprisingly, our sample of children with MLD did not perform poorly on symbolic quantity  
484 discrimination compared to TD children either before ( $t(50) = -.71$ ,  $p = .48$ , Cohen's  $d = -.21$ ,  $BF$   
485 = .36) or after ( $t(50) = -1.01$ ,  $p = .32$ , Cohen's  $d = -.30$ ,  $BF = .44$ ) training, with comparable  
486 variability between groups across time (coefficient of variation range: TD children: .23 – .31;  
487 children with MLD: .26 – .35; test of variance:  $F_s < 1.14$ ,  $p_s > .73$ ). Post-hoc analysis revealed  
488 that children's performance on all measures of WJ-III math subtests were not significantly  
489 correlated with symbolic quantity discrimination before training in either group ( $p_s < .21$ ,  $p_s > .42$ ,  
490  $BFs < .65$ ). These results suggest that there is a large variability in these children's ability to  
491 perform on basic numerical tasks and that mathematical difficulties may be present even in the  
492 absence of number sense deficits.

493  
494 ***Association between intrinsic functional connectivity of the hippocampus and training-***  
495 ***induced number sense learning***

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

503  
504 **Training group.** Functional connectivity of the left and right hippocampal ROIs with the left IPS  
505 before training was positively correlated with number sense learning in the training group (height  
506 threshold  $p < .005$ , cluster extent threshold  $p < .05$ ) (Figures 3A-B; Tables 6-7). Similar results  
507 were observed at a more stringent height threshold ( $p < .001$ , uncorrected). A conjunction  
508 analysis of connectivity patterns across the left and right hippocampal ROIs confirmed a single  
509 overlapping target region in the left IPS, as identified by Juelich Histological Atlas, positively  
510 associated with number sense learning (Figure 3C). The left cuneus was identified as an  
511 overlapping target region negatively associated with number sense learning. Follow-up ROI-  
512 based correlation analyses were conducted in the training and control group, using the  
513 functional connectivity between both the left and right hippocampal regions and the left IPS  
514 region identified in the training group. Similar to results from whole-brain regression analysis,  
515 hippocampal functional connectivity with the left IPS predicted number sense learning in the  
516 training group (left hippocampus:  $\rho = .42$ ,  $p = .002$ , BF = 18.10; right hippocampus:  $\rho = .41$ ,  $p$   
517 = .003, BF = 9.93; Figures 4A-B; Table 8). This association was not significant in the control  
518 group (left hippocampus:  $\rho = -.20$ ,  $p = .45$ , BF = .56; right hippocampus:  $\rho = -.14$ ,  $p = .59$ , BF  
519 = .53). A balanced fourfold cross-validation combined with linear regression (see *Methods*)  
520 further validated the robustness of findings in the training group. Functional connectivity of the  
521 left hippocampus [ $r(pred,actual) = .33$ ,  $p = .002$ ] and the right hippocampus [ $r(pred,actual) = .32$ ,

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

525  
526 Direct comparisons of correlation coefficients between training and control groups revealed a  
527 significant difference in the relationship between functional connectivity with the left IPS and  
528 learning for the left ( $Z = 2.15, p = .02$ ) and the right hippocampus ( $Z = 1.90, p = .03$ ). Additional  
529 correlational analyses were conducted to examine the specificity of left-lateralized IPS functional  
530 connectivity patterns associated with learning. Here, the functional connectivity of the left and  
531 the right hippocampus with the right IPS (a contralateral region of the left IPS identified in the  
532 whole training group) did not significantly relate to number sense learning in the training or  
533 control group ( $ps < .16, ps > .28, BFs < .56$ ; **Table 8**).

534  
535 Additional analysis confirmed that IPS regions identified from the current study overlap with the  
536 left IPS region identified from Neurosynth based meta-analysis, using the term “arithmetic” as  
537 defined in a previous study (Supekar et al., 2021) (**Figures 5A-B**), which indicates that the IPS  
538 region identified from our whole brain analysis converges with the region previously shown to be  
539 involved in math cognition. Finally, when using the IPS region identified from Neurosynth based  
540 meta-analysis, the association between hippocampal-left IPS circuits and learning in response  
541 to number sense training remains significant (left hippocampus-left IPS:  $\rho = .35, p = .011$ ; right  
542 hippocampus-left IPS:  $\rho = .30, p = .03$ ; **Figures 5C-D**).

543  
544 Taken together, these results demonstrate that bilateral hippocampal functional connectivity  
545 with a common target in the left IPS is predictive of learning in response to a 4-week number  
546 sense training.

547

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

567

568 Using the left IPS region identified in the whole training group as target ROI, we found that  
569 hippocampal functional connectivity with the left IPS is positively associated with number sense  
570 learning in the TD group (left hippocampus:  $\rho = .43$ ,  $p = .01$ ,  $BF = 3.98$ ; right hippocampus:  $\rho$   
571  $= .38$ ,  $p = .03$ ,  $BF = 4.58$ ; **Table 9**) as well as in children with MLD (left hippocampus:  $\rho = .52$ ,  $p$   
572  $= .03$ ,  $BF = 2.20$ ; right hippocampus:  $\rho = .52$ ,  $p = .03$ ,  $BF = .93$ ). In a cross-validation analysis  
573 (see *Methods*), functional connectivity of the left [ $r(pred, actual) = .26$ ,  $p = .02$ ] and right

574 hippocampus [ $r(pred,actual) = .37, p = .006$ ] with the left IPS was predictive of learning in TD  
575 children. In children with MLD, the relationship between functional connectivity and learning did  
576 not reach statistical significance at  $p < .05$  in the left [ $r(pred,actual) = .22, p = .06$ ] and right  
577 [ $r(pred,actual) = .10, p = .15$ ] hippocampus.

578  
579 Direct comparisons of correlation coefficients between TD and MLD groups revealed no  
580 significant difference in the relationship between hippocampal functional connectivity with the  
581 left IPS and learning gains (left hippocampus:  $Z = .37, p = .36$ ; right hippocampus:  $Z = .56, p$   
582  $= .29$ ). Finally, similar to the results from whole training group, functional connectivity of the left  
583 and right hippocampus with the right IPS did not significantly relate to learning in TD children or  
584 children with MLD ( $ps < .33, ps > .19, BFs < .55$ ; **Table 9**).

585  
586 To further address whether the relation between hippocampal-parietal functional connectivity  
587 and number sense learning varies as a function of individual differences in math ability, we  
588 additionally used a dimensional approach. In a multiple regression model, number sense  
589 learning (gains in symbolic quantity discrimination task efficiency) was entered as dependent  
590 variable, hippocampal-parietal connectivity (left hippocampus – left IPS or right hippocampus –  
591 left IPS link), math ability (WJ-III Math Fluency), and interaction between hippocampal-parietal  
592 connectivity and math ability were entered as independent variables, and pre-training symbolic  
593 number comparison efficiency was entered as a covariate. We found a significant main effect of  
594 hippocampal-parietal connectivity (left hippocampus – left IPS:  $b = .36, se = .12, t = 3.13, p$   
595  $= .003$ ; right hippocampus – left IPS:  $b = .32, se = .12, t = 2.57, p = .01$ ) but no significant main  
596 effect of math ability or interaction between hippocampal-parietal connectivity and math ability  
597 ( $ts < .81, ps > .42$ ) on number sense learning.

598

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

607

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

610 To determine whether structural integrity plays a similar role in learning as functional circuitry,  
611 we examined the relation between pre-training hippocampal–parietal white matter connectivity  
612 and number sense learning. Using probabilistic tractography of HARDI data, we identified long-  
613 range anatomical connections between the hippocampus and IPS, identified from functional  
614 connectivity analysis, averaged across all children (**Figure 6**; see *Methods* for details) and  
615 assessed the relation between hippocampal–parietal structural connectivity and number sense  
616 learning. In contrast to the functional connectivity results, however, we did not observe evidence  
617 for structural connectivity of the hippocampus with IPS associated with training-related gains in  
618 symbolic quantity discrimination task ( $|\rho|s < .21$ ,  $ps > .18$ ,  $BFs < .62$ ; **Table 8**). Further, there  
619 were no significant associations between structural connectivity between hippocampus and IPS  
620 and number sense learning in TD children and children with MLD ( $|\rho|s < .39$ ,  $ps > .16$ ,  $BFs <$   
621  $1.34$ ; **Table 9**).

622

623 To further address potential contribution of structural connectivity measures to learning, we  
624 conducted multiple regression analysis to determine whether functional and structural

625 connectivity measures together predicted number sense learning better than each measure  
626 alone. Specifically, we examined whether the full model (Model 4) including all hippocampal–  
627 IPS functional and structural connectivity measures as predictors, compared to including  
628 functional or structural connectivity measures alone (Model 2 and Model 3, respectively), better  
629 predict number sense learning (**Table 10**). Here we found that the full model (Model 4)  
630 explained most variance in number sense learning (adjusted  $R^2 = .54$ ,  $F(9,33) = 6.55$ ,  $p < .001$ ),  
631 significantly better than the model including structural connectivity measures alone (Model 3;  
632  $\Delta R^2 = .33$ ,  $p < .001$ ,  $BF > 100$ ). Critically, there was insufficient evidence ( $.33 < BF < 3$ ) that the  
633 full model including both functional and structural connectivity measures (Model 4) explain  
634 additional variance in learning, compared to the model including functional connectivity  
635 measures alone (Model 2;  $\Delta R^2 = .14$ ,  $p = .026$ ,  $BF = 2.26$ ). Thus, we did not observe evidence  
636 that structural connectivity measures jointly predict number sense learning over and above  
637 functional connectivity measures.

638

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

644

#### 645 ***Reverse meta-analysis of associations between hippocampal–parietal circuits and*** 646 ***cognitive functions***

647 To further determine the functional role of hippocampal–IPS circuits identified in the current  
648 study, we conducted a reverse meta-analysis across 14,371 published fMRI studies up to July  
649 2018 from NeuroSynth (Yarkoni et al., 2011) database in relation to 89 cognitive atlas terms  
650 CogAt (Poldrack et al., 2011). To perform reverse meta-analysis relating a cognitive function

651 with a specific circuit, we computed the probability that a term associated with a cognitive  
652 function is mentioned in a study, given that the study jointly reports activations in the left or right  
653 hippocampus and left or right IPS. We considered sufficient evidence for an association if its  
654 probability is amongst the 5% most probable term associations for all analyzed circuits (see  
655 *Methods* for details). This meta-analysis tool allowed us to synthesize a wealth of findings from  
656 previous research and generalize the findings of hippocampal–parietal circuits across various  
657 tasks and analysis approaches (Muller et al., 2018).

658  
659 Our results from reverse meta-analysis show that co-activations of both the left and right  
660 hippocampus and IPS are significantly associated with the term *learning* as well as related  
661 terms, *encoding*, *memory*, and *retrieval* (**Figure 7**). The term *recognition* was associated with  
662 co-activations of the left hippocampus and left IPS and those of bilateral hippocampus and right  
663 IPS. Two terms, *attention* and *working memory*, were associated with co-activations of the right  
664 hippocampus and bilateral IPS. The term *emotion* was associated with co-activations of the right  
665 hippocampus and right IPS. Finally, the term *perception* was associated with co-activations of  
666 the left hippocampus and left IPS. Notably, no other cognitive atlas terms were significantly  
667 associated with hippocampal–parietal functional circuits. These meta-analytic findings from a  
668 large set of fMRI studies expand on findings from our training study and provide converging  
669 evidence for a strong association between hippocampal–parietal functional circuitry and learning  
670 and related functions.

671

## 672 **Discussion**

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

677 left IPS functional circuits predict number sense training gains. Our findings provide important  
678 insights into brain-based biomarkers for early identification of individual differences in  
679 acquisition of number sense and inform interventions targeting individual needs (Hale et al.,  
680 2010).

681  
682 We show that our integrative number sense training is effective across children with a wide  
683 range of abilities. Our results build on previous training studies that enhance understanding of  
684 numerical magnitudes in children (Dyson, Jordan, & Glutting, 2013; Kucian et al., 2011; Wilson  
685 et al., 2006). Our study maximized effectiveness of training by uniquely combining computerized  
686 games with tutoring activities using physical manipulatives and provide new insights into the  
687 development of effective 'hybrid' interventions across children with different backgrounds. More  
688 generally, individualized training programs designed to enhance integration of symbolic and  
689 non-symbolic representations of quantity may have the potential to build strong foundations for  
690 mathematical learning across all children.

691  
692 It is noteworthy that prior to training, we did not observe poor number sense in our sample of  
693 children with MLD, who performed significantly worse than TD children on assessments of  
694 arithmetic problem solving and mathematical reasoning, similar to previous observation that  
695 difficulties in math problem solving may be present even in the absence of number sense  
696 deficits (Peters, Op de Beeck, & De Smedt, 2020). As MLD is considered a heterogeneous  
697 disorder with multiple cognitive deficits (Fias, Menon, & Szucs, 2013; Kaufmann et al., 2013),  
698 further studies that employ assessments in various cognitive domains may help determine  
699 multidimensional neurocognitive deficits in MLD. In addition, development of classification of  
700 subtypes of MLD will be an important avenue for future research.

701

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

712

713 Our finding is also consistent with previous evidence indicating the role of IPS in representation  
714 of quantities (Cohen Kadosh, Cohen Kadosh, Kaas, Henik, & Goebel, 2007; Piazza & Eger,  
715 2016). Remarkably, functional connectivity of the left and right hippocampus identified a single  
716 region in the left IPS that predicted learning. Previous studies have found that compared to its  
717 right hemisphere homolog, the left IPS is particularly important for symbolic number processing  
718 (Ansari, 2007; Bugden, Price, McLean, & Ansari, 2012; Piazza, Pinel, Le Bihan, & Dehaene,  
719 2007; Sokolowski, Fias, Mousa, & Ansari, 2017) and that with age and increased proficiency in  
720 numerical skills, there is an increase in left IPS activity (Ansari, 2008; Bugden, DeWind, &  
721 Brannon, 2016; Emerson & Cantlon, 2015; Rivera, Reiss, Eckert, & Menon, 2005). In this  
722 context, it is possible that our finding of left-lateralized IPS response may be reflective of  
723 increased proficiency for symbolic numbers.

724

725 In addition to the left IPS as a converging target region for the left and right hippocampal  
726 functional circuits positively associated with number sense learning, the left cuneus, implicated  
727 in low-level visual processing (Vanni, Tanskanen, Seppa, Uutela, & Hari, 2001), was identified

728 as a target region negatively associated with number sense training gains. This finding suggests  
729 that children's number sense learning likely relies on enhanced semantic representation of  
730 quantity, rather than visual perception of numbers. In fact, greater hippocampal functional  
731 connectivity with a visual region (cuneus) predicted poor learning. No other brain regions were  
732 identified as overlapping target regions across the left and right hippocampal functional  
733 connectivity associated with learning. Taken together, our findings demonstrate a key role for  
734 hippocampal–parietal circuits in children's number sense learning.

735  
736 Our whole brain analysis of subgroups of children revealed that TD children recapitulate the  
737 hippocampal–left IPS functional circuit-related learning as seen in the combined group. While no  
738 significant target regions were detected in hippocampal connectivity positively associated with  
739 learning in children with MLD possibly due to modest sample size in this group, additional  
740 analysis confirmed that the association between hippocampal–left IPS circuits and learning was  
741 similar in the TD and MLD groups, consistent with our observation of comparable training gains  
742 in the two groups. Thus, our findings identify hippocampal–left IPS functional circuit as a novel  
743 locus of learning that supports acquisition of fundamental building blocks of numerical  
744 proficiency across all children, including those with learning disabilities. Future studies with a  
745 larger sample of children with MLD may further clarify heterogeneous profiles of learning-related  
746 hippocampal circuits.

747  
748 Our probabilistic tractography of HARDI data indicates the presence of long-range anatomical  
749 connections between the hippocampus and the parietal cortex in 7-10 year old children, similar  
750 to the observation in younger children (Ngo et al., 2017). In contrast to findings from functional  
751 circuit analysis, however, structural connectivity between the hippocampus and IPS did not  
752 relate to individual differences in learning. In addition, while hippocampal–parietal functional  
753 connectivity measures jointly predicted learning over and above measures of structural integrity,

754 we did not find evidence that structural connectivity measures jointly predict learning over and  
755 above functional connectivity measures. Thus, although hippocampal–parietal white matter  
756 tracts appear to be formed by early childhood, how they contribute to number sense learning  
757 remains unresolved. Crucially, our study provides evidence for emergent functional properties of  
758 hippocampal–parietal circuits as significant and independent neural predictors of number sense  
759 learning, which may inform early identification of individual differences in response to  
760 intervention. Future studies employing larger data sets with an active control group may help  
761 identify generalizable predictive features of learning and further determine specific mechanisms  
762 underlying acquisition of foundational cognitive skills.

763  
764 Finally, our findings converge on results from a reverse meta-analysis in which we examined the  
765 role of hippocampal–parietal functional links identified in the present study. Across 14,371 fMRI  
766 studies and 89 cognitive atlas terms, our analysis revealed a significant association between  
767 bilateral hippocampal–IPS functional circuits and the term *learning*, along with related terms  
768 *memory*, *encoding*, and *retrieval*. Interactions between the hippocampus and neocortex are  
769 known to be crucial for memory formation (McClelland, McNaughton, & O'Reilly, 1995; Tse et al.,  
770 2007) and hippocampal connectivity with parietal and frontal cortical regions have been shown  
771 to be associated with longitudinal gains in memory retrieval fluency in children (Qin et al., 2014).  
772 Taken together, our findings identify specific hippocampal–neocortical functional circuitries that  
773 may contribute to learning and memory consolidation.

774  
775 While involvement of the hippocampus in learning and memory is well known, the specific role  
776 of hippocampal–IPS functional circuits has been less clear, as research on the role of parietal  
777 cortex in memory has emphasized its angular gyrus subdivision in episodic memory (Sestieri,  
778 Shulman, & Corbetta, 2017). The angular gyrus, as part of the default mode network, is crucial  
779 for generating integrated representation of information retrieved from episodic memory (Binder

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

792

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

800 **References**

- 801 Andersson, J. L. R., & Sotiropoulos, S. N. (2016). An integrated approach to correction for off-  
802 resonance effects and subject movement in diffusion MR imaging. *NeuroImage*, *125*,  
803 1063-1078. doi:10.1016/j.neuroimage.2015.10.019
- 804 Ansari, D. (2007). Does the parietal cortex distinguish between "10," "ten," and ten dots?  
805 *Neuron*, *53*(2), 165-167. doi:10.1016/j.neuron.2007.01.001
- 806 Ansari, D. (2008). Effects of development and enculturation on number representation in the  
807 brain. *Nature reviews neuroscience*, *9*, 278-291. doi:10.1038/hnr2334
- 808 Avants, B. B., Schoenemann, P. T., & Gee, J. C. (2006). Lagrangian frame diffeomorphic image  
809 registration: Morphometric comparison of human and chimpanzee cortex. *Med Image*  
810 *Anal*, *10*(3), 397-412. doi:10.1016/j.media.2005.03.005
- 811 Barnett, A. G., van der Pols, J. C., & Dobson, A. J. (2005). Regression to the mean: What it is  
812 and how to deal with it. *International Journal of Epidemiology*, *34*, 215–220.
- 813 Behrens, T. E., Berg, H. J., Jbabdi, S., Rushworth, M. F., & Woolrich, M. W. (2007). Probabilistic  
814 diffusion tractography with multiple fibre orientations: What can we gain? *NeuroImage*,  
815 *34*(1), 144-155. doi:10.1016/j.neuroimage.2006.09.018
- 816 Binder, J. R., & Desai, R. H. (2011). The neurobiology of semantic memory. *Trends Cogn Sci*,  
817 *15*(11), 527-536. doi:10.1016/j.tics.2011.10.001
- 818 Blair, K. P. (2013a). *Feedback in Critter Corral: The effectiveness of implication versus*  
819 *corrective feedback in a math learning game*. Paper presented at the 2013 Early  
820 Education and Technology for Children Conference, Salt Lake City, UT.
- 821 Blair, K. P. (2013b). *Learning in Critter Corral: Evaluating three kinds of feedback in a preschool*  
822 *math game*. Paper presented at the Proceedings of the Interaction Design and Children  
823 2013 Conference, New York: ACM.

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

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

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

832 Carter, C. S., Lesh, T. A., & Barch, D. M. (2016). Thresholds, power, and sample sizes in  
833 clinical neuroimaging. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*,  
834 1, 99–100.

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

839 Cho, S., Metcalfe, A. W. S., Young, C. B., Ryali, S., Geary, D. C., & Menon, V. (2012).  
840 Hippocampal–Prefrontal Engagement and Dynamic Causal Interactions in the  
841 Maturation of Children's Fact Retrieval. *Journal of Cognitive Neuroscience*, 24, 1849-  
842 1866. doi:10.1162/jocn\_a\_00246

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

846 Cohen, J. R., Asarnow, R. F., Sabb, F. W., Bilder, R. M., Bookheimer, S. Y., Knowlton, B. J., &  
847 Poldrack, R. A. (2010). Decoding developmental differences and individual variability in  
848 response inhibition through predictive analyses across individuals. *Front Hum Neurosci*,  
849 4, 47. doi:10.3389/fnhum.2010.00047

850 Cohen Kadosh, R., Cohen Kadosh, K., Kaas, A., Henik, A., & Goebel, R. (2007). Notation-  
851 dependent and -independent representations of numbers in the parietal lobes. *Neuron*,  
852 53(2), 307-314. doi:10.1016/j.neuron.2006.12.025

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

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

860 Desikan, R., Ségonne, F., Fischl, B., Quinn, B., Dickerson, B., Blacker, D., . . . Killiany, R.  
861 (2006). An automated labeling system for subdividing the human cerebral cortex on MRI  
862 scans into gyral based regions of interest. *NeuroImage*, 31, 968-980.

863 Donahue, C. J., Sotiropoulos, S. N., Jbabdi, S., Hernandez-Fernandez, M., Behrens, T. E.,  
864 Dyrby, T. B., . . . Glasser, M. F. (2016). Using Diffusion Tractography to Predict Cortical  
865 Connection Strength and Distance: A Quantitative Comparison with Tracers in the  
866 Monkey. *The Journal of Neuroscience*, 36(25), 6758-6770. doi:10.1523/jneurosci.0493-  
867 16.2016

868 Dyson, N. I., Jordan, N. C., & Glutting, J. (2013). A Number Sense Intervention for Low-Income  
869 Kindergartners at Risk for Mathematics Difficulties. . *Journal of Learning Disabilities*, 46  
870 166–181.

871 Eichenbaum, H. (2004). Hippocampus: Cognitive processes and neural representations that  
872 underlie declarative memory. *Neuron*, 44(1), 109-120. doi:DOI  
873 10.1016/j.neuron.2004.08.028

874 Eickhoff, S. B., Stephan, K. E., Mohlberg, H., Grefkes, C., Fink, G. R., Amunts, K., & Zilles, K.  
875 (2005). A new SPM toolbox for combining probabilistic cytoarchitectonic maps and  
876 functional imaging data. *NeuroImage*, *25*, 1325-1335.

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

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

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

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

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

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

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

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

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

903 Iuculano, T., Rosenberg-Lee, M., Richardson, J., Tenison, C., Fuchs, L., Supekar, K., & Menon,  
904 V. (2015). Cognitive tutoring induces widespread neuroplasticity and remediates brain  
905 function in children with mathematical learning disabilities. *Nature Communications*, *6*,  
906 8453. doi:10.1038/ncomms9453

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

911 Jolles, D., Ashkenazi, S., Kochalka, J., Evans, T., Richardson, J., Rosenberg-Lee, M., . . .  
912 Menon, V. (2016). Parietal hyper-connectivity, aberrant brain organization, and circuit-  
913 based biomarkers in children with mathematical disabilities. *Developmental science*, *19*,  
914 613-631. doi:10.1111/desc.12399

915 Jordan, N. C., Kaplan, D., Ramineni, C., & Locuniak, M. N. (2009). Early math matters:  
916 kindergarten number competence and later mathematics outcomes. *Dev Psychol*, *45*(3),  
917 850-867. doi:10.1037/a0014939

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

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

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

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

928 Kucian, K., Grond, U., Rotzer, S., Henzi, B., Schonmann, C., Plangger, F., . . . von Aster, M.  
929 (2011). Mental number line training in children with developmental dyscalculia.  
930 *NeuroImage*, *57*(3), 782-795. doi:10.1016/j.neuroimage.2011.01.070

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

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

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

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

942 Muller, V. I., Cieslik, E. C., Laird, A. R., Fox, P. T., Radua, J., Mataix-Cols, D., . . . Eickhoff, S. B.  
943 (2018). Ten simple rules for neuroimaging meta-analysis. *Neurosci Biobehav Rev*, *84*,  
944 151-161. doi:10.1016/j.neubiorev.2017.11.012

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

947 Ngo, C. T., Alm, K. H., Metoki, A., Hampton, W., Riggins, T., Newcombe, N. S., & Olson, I. R.  
948 (2017). White matter structural connectivity and episodic memory in early childhood. *Dev*  
949 *Cogn Neurosci*, *28*, 41-53. doi:10.1016/j.dcn.2017.11.001

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

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

955 Peters, L., Op de Beeck, H., & De Smedt, B. (2020). Cognitive correlates of dyslexia,  
956 dyscalculia and comorbid dyslexia/dyscalculia: Effects of numerical magnitude  
957 processing and phonological processing. *Research in Developmental Disabilities*,  
958 *107*(March), 103806-103806. doi:10.1016/j.ridd.2020.103806

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

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

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

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

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

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

976 Rivera, S. M., Reiss, a. L., Eckert, M. a., & Menon, V. (2005). Developmental changes in mental  
977 arithmetic: evidence for increased functional specialization in the left inferior parietal

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

980 Roediger, H. L., Agarwal, P. K., McDaniel, M. A., & McDermott, K. B. (2011). Test-enhanced  
981 learning in the classroom: long-term improvements from quizzing. *J Exp Psychol Appl*,  
982 17(4), 382-395. doi:10.1037/a0026252

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

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

991 Rosenberg-Lee, M., Luculano, T., Bae, S. R., Richardson, J., Qin, S., Jolles, D., & Menon, V.  
992 (2018). Short-term cognitive training recapitulates hippocampal functional changes  
993 associated with one year of longitudinal skill development. *Trends in Neuroscience and*  
994 *Education*. doi:10.1016/j.tine.2017.12.001

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

998 Sestieri, C., Shulman, G. L., & Corbetta, M. (2017). The contribution of the human posterior  
999 parietal cortex to episodic memory. *Nat Rev Neurosci*, 18(3), 183-192.  
1000 doi:10.1038/nrn.2017.6

1001 Sokolowski, H. M., Fias, W., Mousa, A., & Ansari, D. (2017). Common and distinct brain regions  
1002 in both parietal and frontal cortex support symbolic and nonsymbolic number processing

1003 in humans: A functional neuroimaging meta-analysis. *NeuroImage*, 146, 376-394.  
1004 doi:10.1016/j.neuroimage.2016.10.028

1005 Supekar, K., Chang, H., Mistry, P. K., Luculano, T., & Menon, V. (2021). Cognitive training-  
1006 induced modular reorganization of hippocampal-cortical circuits drives learning gains  
1007 and efficient strategies. *Communications Biology*, 4.

1008 Supekar, K., Swigart, A. G., Tenison, C., Jolles, D. D., Rosenberg-lee, M., Fuchs, L., & Menon,  
1009 V. (2013). Neural predictors of individual differences in response to math tutoring in  
1010 primary-grade school children. *Proc Natl Acad Sci*, 110, 8230-8235.

1011 Thomas, C., Ye, F. Q., Irfanoglu, M. O., Modi, P., Saleem, K. S., Leopold, D. A., & Pierpaoli, C.  
1012 (2014). Anatomical accuracy of brain connections derived from diffusion MRI  
1013 tractography is inherently limited. *Proc Natl Acad Sci U S A*, 111(46), 16574-16579.  
1014 doi:10.1073/pnas.1405672111

1015 Thompson, D. D., Lingsma, H. F., Whiteley, W. N., Murray, G. D., & Steyerberg, E. W. (2015).  
1016 Covariate adjustment had similar benefits in small and large randomized controlled trials.  
1017 *Journal of Clinical Epidemiology*, 68, 1068–1075.

1018 Townsend, J. T., & Ashby, F. G. (1978). *Methods of modeling capacity in simple processing*  
1019 *systems* (Vol. 3). Erlbaum, Hillsdale, N.J.: Cognitive Theory.

1020 Tse, D., Langston, R. F., Kakeyama, M., Bethus, I., Spooner, P. A., Wood, E. R., . . . Morris, R.  
1021 G. M. (2007). Schemas and Memory Consolidation. *Science*, 316, 76–82.

1022 Tuch, D. S., Reese, T. G., Wiegell, M. R., Makris, N., Belliveau, J. W., & Wedeen, V. J. (2002).  
1023 High angular resolution diffusion imaging reveals intravoxel white matter fiber  
1024 heterogeneity. *Magnetic Resonance in Medicine: An Official Journal of the International*  
1025 *Society for Magnetic Resonance in Medicine*, 48, 577-582.

1026 Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., . . .  
1027 Joliot, M. (2002). Automated anatomical labeling of activations in SPM using a

1028 macroscopic anatomical parcellation of the MNI MRI single-subject brain. *NeuroImage*,  
1029 15 273-289.

1030 Uddin, L. Q., Supekar, K., Amin, H., Rykhlevskaia, E., Nguyen, D. A., Greicius, M. D., & Menon,  
1031 V. (2010). Dissociable connectivity within human angular gyrus and intraparietal sulcus:  
1032 evidence from functional and structural connectivity. *Cereb Cortex*, 20(11), 2636-2646.  
1033 doi:10.1093/cercor/bhq011

1034 Vanni, S., Tanskanen, T., Seppa, M., Uutela, K., & Hari, R. (2001). Coinciding early activation of  
1035 the human primary visual cortex and anteromedial cuneus. *Proc Natl Acad Sci U S A*,  
1036 98(5), 2776-2780. doi:10.1073/pnas.041600898

1037 Ward, B. D. (2000). Simultaneous Inference for FMRI Data. In *AFNI 3dDeconvolve*  
1038 *Documentation*. Milwaukee, WI: Medical College of Wisconsin.

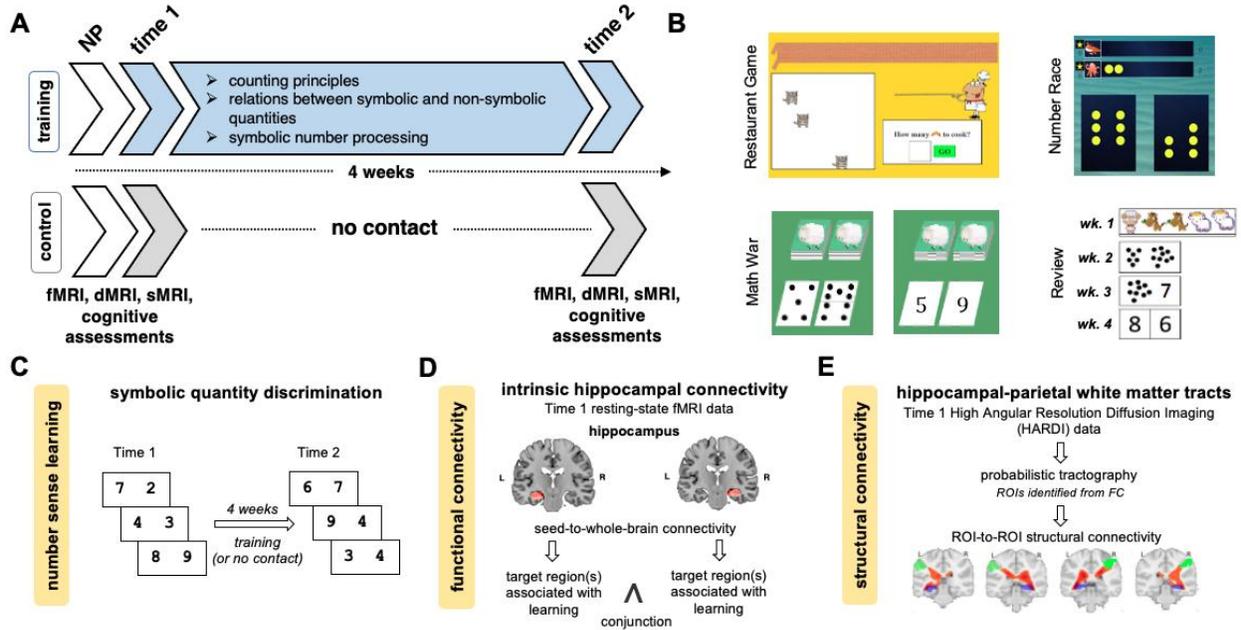
1039 Wechsler, D. (1999). *The Wechsler Abbreviated Scale of Intelligence*. San Antonio, TX: The  
1040 Psychological Corp.

1041 Wilson, A. J., Revkin, S. K., Cohen, D., Cohen, L., & Dehaene, S. (2006). An open trial  
1042 assessment of "the number race", an adaptive computer game for remediation of  
1043 dyscalculia. *Behavioral and Brain Functions*, 2, 20. doi:10.1186/1744-9081-2-20

1044 Woodcock, R. W., McGrew, K. S., & Mather, N. (2001). *Woodcock-Johnson III Tests of*  
1045 *Achievement*. Istatasca, IL: Riverside.

1046 Yarkoni, T., Poldrack, R. a., Nichols, T. E., Van Essen, D. C., & Wager, T. D. (2011). Large-  
1047 scale automated synthesis of human functional neuroimaging data. *Nature methods*, 8,  
1048 665-670. doi:10.1038/nmeth.1635

1049 Zeithamova, D., & Bowman, C. R. (2020). Generalization and the hippocampus: More than one  
1050 story? *Neurobiology of Learning and Memory*, 175, 107317 Contents.



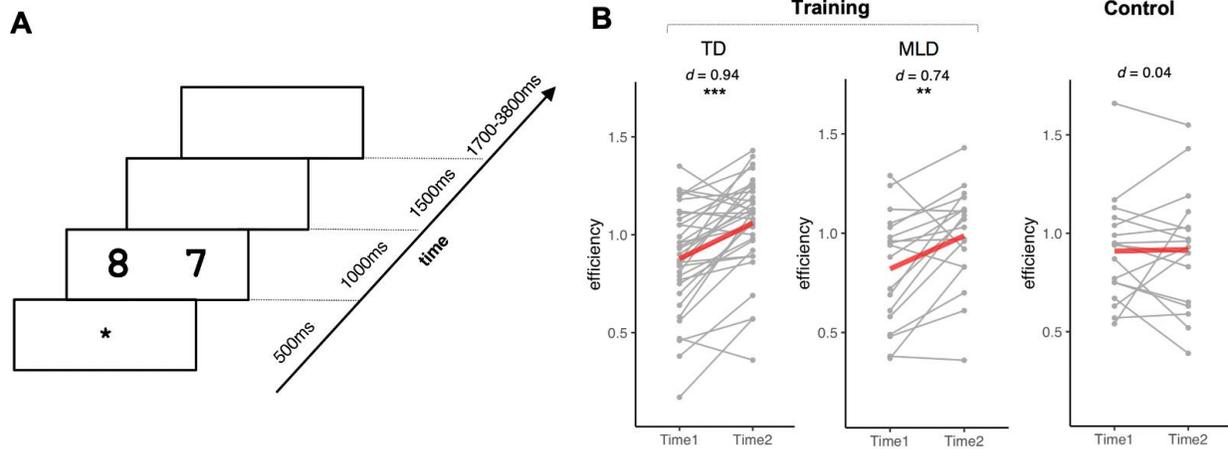
1052 **Figure 1. Overview of study design, sample training materials, and schematic illustration**  
 1053 **of analysis approach.**

1054 **A. Overview of study design.** The study involved multiple sessions: pre-training demographic  
 1055 and neuropsychological (NP) assessments; pre-training (time 1) cognitive assessments and  
 1056 brain imaging, including task-related and resting-state functional MRI (fMRI), diffusion MRI  
 1057 (dMRI) using a High Angular Resolution Diffusion Imaging sequence, and high-resolution  
 1058 structural MRI (sMRI); number sense training (in the training group; see *Methods* for details) or  
 1059 no contact (in the control group); and post-training (time 2) brain imaging and cognitive  
 1060 assessments identical to pre-training. Children in the training group engaged in progressive  
 1061 learning activities across four weeks to strengthen their understanding of the relations between  
 1062 symbolic and non-symbolic representations of quantity. **B. Sample training materials.** Across 4  
 1063 weeks of one-on-one tutoring sessions, children in the training group completed a variety of  
 1064 activities with a tutor (see *Methods* for details). **C. Number sense learning.** Children’s number  
 1065 sense learning was measured by changes in efficiency in symbolic quantity discrimination task

1066 from time 1 to time 2 in response to 4 weeks of number sense training (or no contact). **D.**

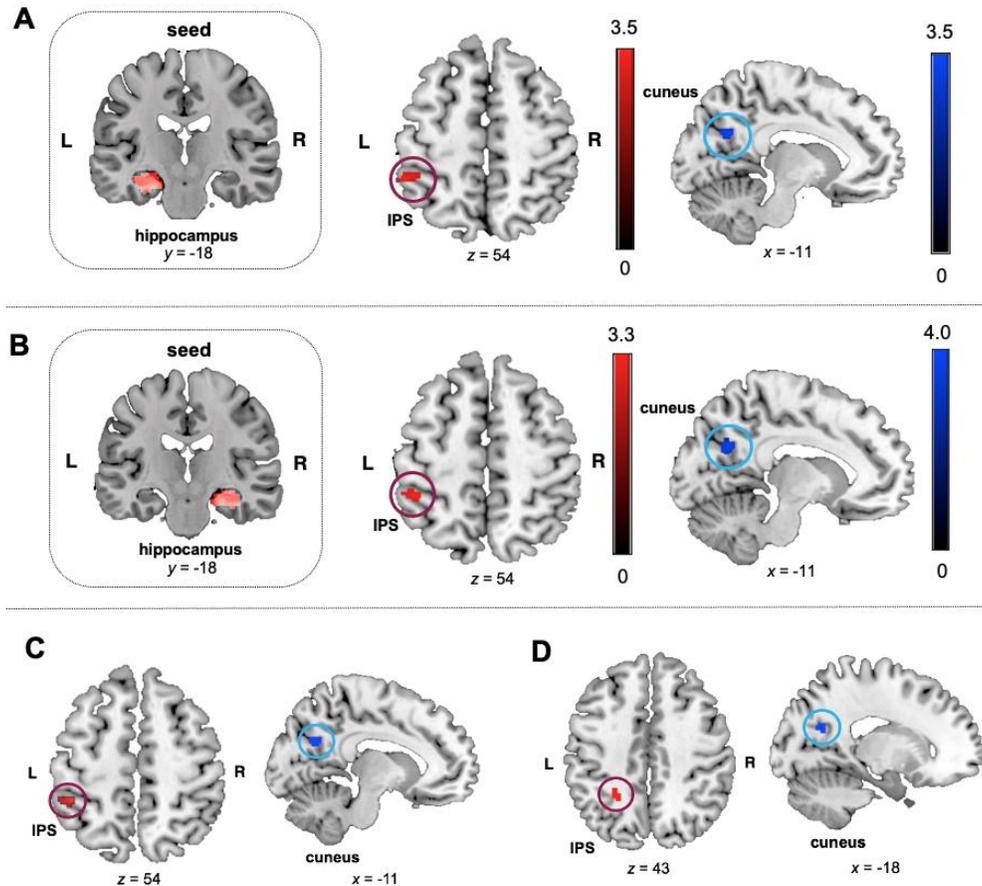
1067 **Functional connectivity.** Using Time 1 resting-state fMRI data, hippocampal seed-to-whole-  
1068 brain connectivity analysis was performed to assess intrinsic functional connectivity of the  
1069 hippocampus predictive of number sense learning. A conjunction analysis was performed  
1070 between left and right hippocampal functional connectivity patterns to identify overlapping target  
1071 regions associated with learning. Region of interest (ROI) based analysis was used to examine  
1072 the association between functional connectivity and learning in training and control groups. **E.**

1073 **Structural connectivity.** Using High Angular Resolution Diffusion Imaging data, probabilistic  
1074 tractography was performed using ROIs identified from functional connectivity (FC) analysis.  
1075 Structural connectivity strengths between the ROIs were estimated for each subject to test the  
1076 associations with number sense learning. ROI-based analysis was performed to examine the  
1077 relation between structural connectivity and learning in training and control groups.



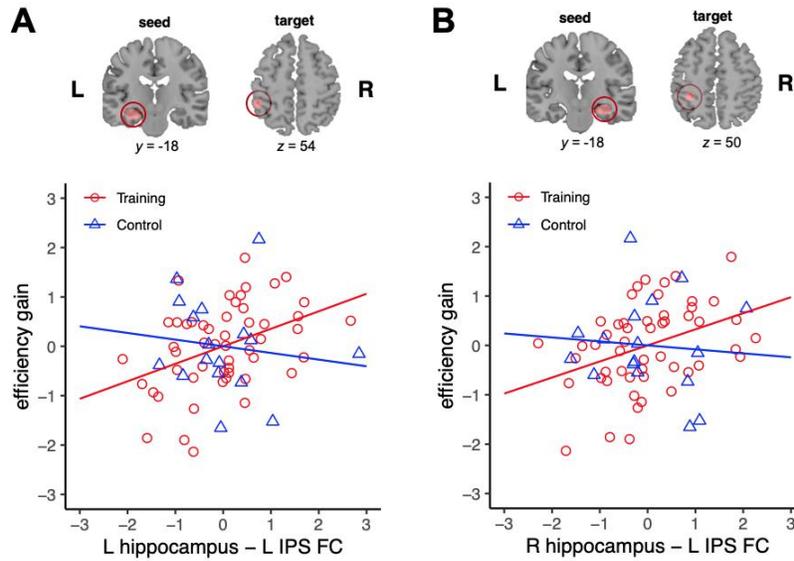
1078 **Figure 2. Symbolic quantity discrimination task design and behavioral results.**

1079 **A. Task design.** Before and after training (or no contact), children completed one run of  
 1080 symbolic quantity discrimination task in the fMRI scanner. Participants first saw a fixation cross,  
 1081 followed by horizontal presentation of two quantities in Arabic numbers. Participants were  
 1082 instructed to press the left button if the left side had a larger quantity and the right side if the  
 1083 right side had a larger quantity. Upon the button press, a blank screen was presented to fill up  
 1084 the response phase, followed by a jittered intertrial interval. The duration of each phase in  
 1085 seconds is reported in parenthesis. **B. Behavioral results.** Four weeks of training improved  
 1086 performance on symbolic quantity discrimination task in both training groups of typically  
 1087 developing (TD) children and children with mathematical learning difficulties (MLD) but not in  
 1088 no-contact control group. \*\*  $p < .01$ ; \*\*\*  $p < .001$ .

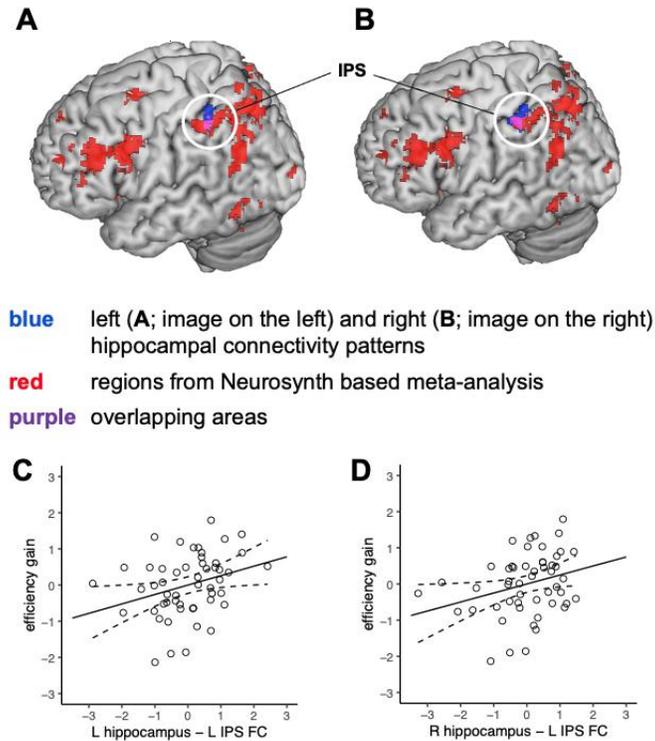


1089 **Figure 3. Hippocampal functional connectivity predicts number sense training gains.**  
 1090 **A, B.** Functional connectivity before training of the left and right hippocampal regions of interest  
 1091 (ROIs; selected from Brainnetome (Fan et al., 2016) parcellations) with the left intraparietal  
 1092 sulcus (IPS) positively (red) predicted learning (efficiency gains in symbolic quantity  
 1093 discrimination) in the training group. Hippocampal connectivity with the left cuneus was  
 1094 negatively (blue) correlated with learning. **C.** Conjunction analysis of functional connectivity  
 1095 patterns revealed the left IPS and the left cuneus to be an overlapping target region across the  
 1096 left and right hippocampal seed-to-whole brain connectivity positively (red) and negatively (blue)  
 1097 associated with learning, respectively, in the training group. **D.** In typically developing (TD)  
 1098 children in the training group, hippocampal functional connectivity with the left IPS was  
 1099 positively associated with learning (red) and that with the left cuneus was negatively associated  
 1100 with learning (blue). In children with mathematical learning difficulties (MLD) in the training

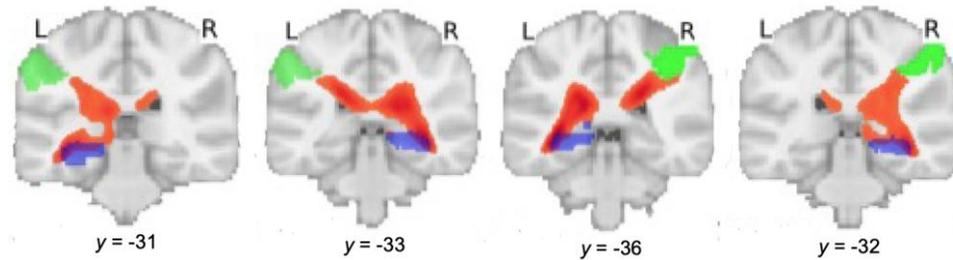
1101 group, conjunction analysis did not yield any overlapping target region across the left and right  
1102 hippocampal seed-to-whole-brain functional connectivity associated with learning. L = left, R =  
1103 right.



1104 **Figure 4. Hippocampal–parietal functional circuits predict number sense training gains.**  
 1105 Scatter plots of the relationship between functional connectivity (FC) of the left (A) and right (B)  
 1106 hippocampus with the left intraparietal sulcus (IPS) and changes in efficiency in symbolic  
 1107 quantity discrimination (efficiency gain) in children who received training (Training) and no-  
 1108 contact control group (Control). Greater FC with the left IPS predicts efficiency gain in the  
 1109 Training group ( $p_s > .40$ ,  $p_s < .004$ ), but not in the Control group ( $|p_s| < .21$ ,  $p_s > .44$ ). Target  
 1110 (IPS) regions of interest were identified using a 6-mm sphere centered on the peak voxel to  
 1111 estimate FC. L = left, R = right.

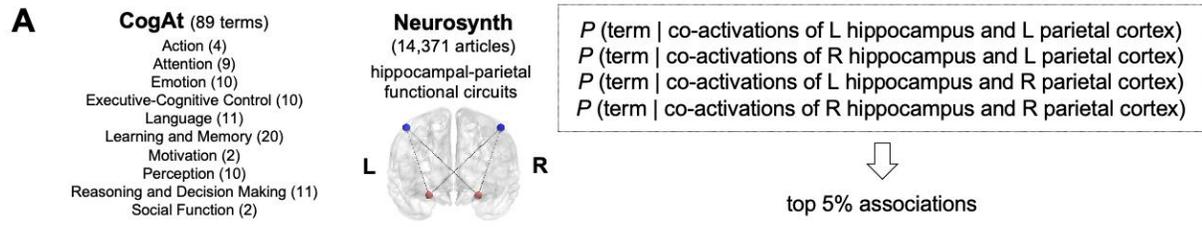


1112 **Figure 5. Hippocampal connectivity with a left intraparietal sulcus (IPS) region, identified**  
 1113 **using Neurosynth based meta-analysis, predicts number sense training gains.** Left IPS  
 1114 regions (in white circle) overlap across (**A**) left and (**B**) right hippocampal connectivity patterns  
 1115 identified in the current study and the IPS region identified from Neurosynth based meta-  
 1116 analysis. (**C-D**). Scatter plots of the relationship between functional connectivity (FC) of the (**A**)  
 1117 left and (**B**) right hippocampal with the left IPS region identified from Neurosynth based meta-  
 1118 analysis. FC between the hippocampus and left IPS correlates with changes in efficiency in  
 1119 symbolic quantity discrimination (efficiency gain) following number sense training ( $p > .29$ ,  $p < .04$ ).  
 1120 The IPS region of interest was identified using a 6-mm sphere centered on the peak  
 1121 voxel to estimate FC. L = left, R = right.



1122 **Figure 6. Hippocampal–parietal white matter tracts in children.**

1123 White matter tracts (depicted in orange) are identified between the hippocampus (purple) and  
 1124 intraparietal sulcus (IPS) (light green), averaged across all children using probabilistic  
 1125 tractography of High Angular Resolution Diffusion Imaging data. IPS regions of interest (ROIs)  
 1126 were selected from Brainnetome (Fan et al., 2016) parcellations that overlapped with target  
 1127 ROIs identified in functional connectivity analysis (left IPS) and contralateral regions (right IPS)  
 1128 to estimate hippocampal–parietal structural connectivity. L = left, R = right.



**B**

<p><u>Action</u></p> <ul style="list-style-type: none"> <li>action</li> <li>motor control</li> <li>movement</li> <li>response selection</li> </ul> <p><u>Attention</u></p> <ul style="list-style-type: none"> <li>attention</li> <li>distraction</li> <li>fixation</li> <li>focus</li> <li>search</li> <li>selective attention</li> <li>spatial attention</li> <li>sustained attention</li> <li>visual attention</li> </ul> <p><u>Emotion</u></p> <ul style="list-style-type: none"> <li>anxiety</li> <li>arousal</li> <li>emotion</li> <li>empathy</li> <li>facial expression</li> <li>fear</li> <li>mood</li> <li>pain</li> <li>stress</li> <li>valence</li> </ul> <p><u>Executive-Cognitive Control</u></p> <ul style="list-style-type: none"> <li>cognitive control</li> </ul>	<p>goal</p> <ul style="list-style-type: none"> <li>inhibition</li> <li>maintenance</li> <li>manipulation</li> <li>monitoring</li> <li>planning</li> <li>response inhibition</li> <li>updating</li> <li>working memory</li> </ul> <p><u>Language</u></p> <ul style="list-style-type: none"> <li>language</li> <li>language comprehension</li> <li>meaning</li> <li>morphology</li> <li>naming</li> <li>semantic memory</li> <li>sentence comprehension</li> <li>speech perception</li> <li>speech production</li> <li>verbal fluency</li> <li>word recognition</li> </ul> <p><u>Learning and Memory</u></p> <ul style="list-style-type: none"> <li>adaptation</li> <li>autobiographical memory</li> <li>consolidation</li> <li>encoding</li> <li>episodic memory</li> <li>expertise</li> </ul>	<p>face recognition</p> <ul style="list-style-type: none"> <li>familiarity</li> <li>knowledge</li> <li>learning</li> <li>memory</li> <li>memory retrieval</li> <li>priming</li> <li>recall</li> <li>recognition</li> <li>rehearsal</li> <li>reinforcement learning</li> <li>retention</li> <li>retrieval</li> <li>skill</li> </ul> <p><u>Motivation</u></p> <ul style="list-style-type: none"> <li>expectancy</li> <li>task difficulty</li> </ul> <p><u>Perception</u></p> <ul style="list-style-type: none"> <li>detection</li> <li>discrimination</li> <li>localization</li> <li>mental imagery</li> <li>multisensory</li> <li>navigation</li> <li>object recognition</li> <li>perception</li> <li>rhythm</li> <li>visual perception</li> </ul>	<p><u>Reasoning and Decision Making</u></p> <ul style="list-style-type: none"> <li>categorization</li> <li>decision</li> <li>decision making</li> <li>inference</li> <li>insight</li> <li>intelligence</li> <li>judgment</li> <li>reasoning</li> <li>risk</li> <li>uncertainty</li> <li>utility</li> </ul> <p><u>Social Function</u></p> <ul style="list-style-type: none"> <li>communication</li> <li>social cognition</li> </ul> <p style="text-align: center;"><b>Top 5% associations</b></p> <ul style="list-style-type: none"> <li>● L HIPP – L IPS</li> <li>● R HIPP – L IPS</li> <li>● L HIPP – R IPS</li> <li>● R HIPP – R IPS</li> </ul>
--	---	---	---

1129 **Figure 7. Reverse meta-analysis of 14,371 fMRI studies and cognitive functions reveals a**  
 1130 **significant association between hippocampal–parietal functional circuits and learning.**  
 1131 **A.** A reverse meta-analysis was performed to map hippocampal–parietal functional circuits  
 1132 identified in the current study to cognitive functions (see *Methods* for details). **B.** Top 5%  
 1133 cognitive functions that are mentioned in published articles where co-activations of the left or  
 1134 right hippocampus (HIPP) and the left or right intraparietal cortex (IPS) are reported. L = left, R  
 1135 = right.

1136 **Table 1.** Training activities in each session (3 sessions/week).

	Lessons	Games	Review
Week 1	<ul style="list-style-type: none"> <li>• Counting</li> <li>• Video of a sock puppet counting</li> </ul>	<ul style="list-style-type: none"> <li>• Restaurant Game (Blair, 2013a, 2013b)</li> </ul>	<ul style="list-style-type: none"> <li>• Counting</li> </ul>
Week 2	<ul style="list-style-type: none"> <li>• Comparison</li> <li>• Math Circles</li> </ul>	<ul style="list-style-type: none"> <li>• Number Race (Wilson et al., 2006)</li> <li>• Math War (Iuculano et al., 2015)</li> <li>• Comparing Speed</li> </ul>	<ul style="list-style-type: none"> <li>• Comparison between non-symbolic quantities</li> </ul>
Week 3	<ul style="list-style-type: none"> <li>• Comparison</li> <li>• Math Circles</li> </ul>	<ul style="list-style-type: none"> <li>• Number Race (Wilson et al., 2006)</li> <li>• Math War (Iuculano et al., 2015)</li> <li>• Comparing Speed</li> </ul>	<ul style="list-style-type: none"> <li>• Comparison between non-symbolic and symbolic quantities</li> </ul>
Week 4	<ul style="list-style-type: none"> <li>• Comparison</li> <li>• Beat Your Score (Chang et al., 2019)</li> </ul>	<ul style="list-style-type: none"> <li>• Number Race (Wilson et al., 2006)</li> <li>• Math War (Iuculano et al., 2015)</li> <li>• Comparing Speed</li> </ul>	<ul style="list-style-type: none"> <li>• Comparison between symbolic quantities</li> </ul>

1137 **Table 2. Resting state fMRI data analysis sample.** Time 1 demographics and  
 1138 neuropsychological measures of 4-week number sense training (Training) and no-contact  
 1139 control (Control) groups.

	Training	Control	$\chi^2$ test or two-sample <i>t</i> -test			
	<i>M</i> ( <i>SD</i> )	<i>M</i> ( <i>SD</i> )	$\chi^2_1$ or <i>t</i> (67)	<i>p</i>	$\varphi$ or Cohen's <i>d</i>	BF
Female to Male ratio	27 : 25	11 : 6	.41	.523	.08	.38
Age	8.21 (.61)	8.32 (.56)	.66	.513	.18	.33
<i>WASI:</i>						
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
<i>WJ-III:</i>						
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.

1140 **Table 3. Resting state fMRI data analysis sample.** Time 1 demographics neuropsychological  
 1141 measures of typically developing (TD) children and children with mathematical learning  
 1142 difficulties (MLD) in the training group (TD\_Training and MLD\_Training).

	TD_Training	MLD_Training	$\chi^2$ test or two-sample <i>t</i> -test			
	<i>M</i> ( <i>SD</i> )	<i>M</i> ( <i>SD</i> )	$\chi^2_1$ or <i>t</i> (50)	<i>p</i>	$\varphi$ or Cohen's <i>d</i>	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
<i>WASI:</i>						
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
<i>WJ-III:</i>						
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.

1143 **Table 4. Diffusion MRI data analysis sample.** Time 1 demographics and neuropsychological  
 1144 measures of 4-week number sense training (Training) and no-contact control (Control) groups.

	Training	Control	$\chi^2$ test or two-sample <i>t</i> -test			BF
	<i>M</i> ( <i>SD</i> )	<i>M</i> ( <i>SD</i> )	$\chi^2_1$ or <i>t</i> (54)	<i>p</i>	$\phi$ or Cohen's <i>d</i>	
Female to Male ratio	22 : 21	7 : 6	<.01	>.999	<.01	.28
Age	8.21 (.64)	8.39 (.56)	.92	.360	.29	.43
<i>WASI:</i>						
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
<i>WJ-III:</i>						
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.

1145 **Table 5. Diffusion MRI data analysis sample.** Time 1 demographics neuropsychological  
 1146 measures of typically developing (TD) children and children with mathematical learning  
 1147 difficulties (MLD) in the training group (TD\_Training and MLD\_Training).

	TD_Training	MLD_Training	$\chi^2$ test or two-sample <i>t</i> -test			
	<i>M</i> ( <i>SD</i> )	<i>M</i> ( <i>SD</i> )	$\chi^2_1$ or <i>t</i> (41)	<i>p</i>	$\varphi$ or Cohen's <i>d</i>	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
<i>WASI:</i>						
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
<i>WJ-III:</i>						
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.

1148 **Table 6.** Brain regions showing positive and negative relation between functional connectivity  
 1149 with the left hippocampus and symbolic quantity discrimination efficiency gain in response to 4-  
 1150 week number sense training.

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

1151 **Table 7.** Brain regions showing positive and negative relation between functional connectivity  
 1152 with the right hippocampus and symbolic quantity discrimination efficiency gain in response to 4-  
 1153 week number sense training.

Region	Number of voxels	Peak intensity	MNI coordinates: x,y,z (mm)
<b>Positive relation</b>			
<b>Training</b>			
L IPS/SPL/SMG	156	3.33	-34 -40 50
<b>Training: TD children</b>			
L IPS/SPL/SMG	75	4.20	-24 -48 44
<b>Training: children with MLD</b>			
(none)			
<b>Negative relation</b>			
<b>Training</b>			
L CUN/PCUN	140	3.92	-12 -62 28
<b>Training: TD children</b>			
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
<b>Training: children with MLD</b>			
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.

1154 **Table 8.** Correlations between functional and structural connectivity and symbolic quantity  
 1155 discrimination efficiency gain in 4-week number sense training (Training) and no-contact control  
 1156 (Control) groups.

	Training ( $N = 52^{\dagger}; 43^{\ddagger}$ )			Control ( $N = 17^{\dagger}; 13^{\ddagger}$ )		
	$\rho$	$p$	BF	$\rho$	$p$	BF
<b>Functional connectivity</b>						
L HIPP - L IPS	.42**	.002	<b>18.10</b>	-.20	.445	.57
R HIPP - L IPS	.41**	.003	<b>9.93</b>	-.14	.589	.53
L HIPP - R IPS	-.10	.500	.33	-.14	.586	.54
R HIPP - R IPS	.15	.289	.34	-.10	.708	.55
<b>Structural connectivity</b>						
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:* <sup>†</sup>number of participants included in functional connectivity analysis. <sup>‡</sup>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  $H_0$ ) (Keyesers et al., 2020).

1157 **Table 9.** Correlations between functional and structural connectivity and symbolic quantity  
 1158 discrimination efficiency gain in typically developing (TD) children and children with  
 1159 mathematical learning difficulties (MLD) in the training group (TD\_Training and MLD\_Training).

	TD_Training ( $N = 34^{\dagger}; 28^{\ddagger}$ )			MLD_Training ( $N = 18^{\dagger}; 15^{\ddagger}$ )		
	$\rho$	$p$	BF	$\rho$	$p$	BF
<b>Functional connectivity</b>						
L HIPP - L IPS	.43*	.012	<b>3.98</b>	.52*	.026	2.20
R HIPP - L IPS	.38*	.025	<b>4.58</b>	.52*	.027	.93
L HIPP - R IPS	-.16	.358	.44	<.01	.997	.50
R HIPP - R IPS	.03	.855	.39	.32	.197	.54
<b>Structural connectivity</b>						
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:* <sup>†</sup>number of participants included in functional connectivity analysis. <sup>‡</sup>number of participants included in structural connectivity analysis. \*  $p < .05$ . Boded 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  $H_0$ ) (Keyesers et al., 2020).

1160 **Table 10.** Model comparison between multiple regression analysis including functional and/or  
 1161 structural connectivity measures as predictors of symbolic quantity discrimination efficiency gain  
 1162 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	$\Delta R^2$	$F$	$df$	$p$	$BF$
Model 2 (Model 1 + FC) vs. Model 1 (Baseline)	.24	4.45**	4	.005	<b>7.33</b>
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	<b>&gt;100</b>

*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  $H_0$ ). BF values below .33 provide evidence of absence (evidence for  $H_0$ ) (Keysers et al., 2020).