



RESEARCH ARTICLE

Short-term number sense training recapitulates long-term neurodevelopmental changes from childhood to adolescence

Yunji Park¹  | Yuan Zhang¹  | Hyesang Chang¹  | Vinod Menon^{1,2,3,4}

¹Department of Psychiatry & Behavioral Sciences, Stanford University School of Medicine, Stanford, California, USA

²Department of Neurology and Neurological Sciences, Stanford University School of Medicine, Stanford, California, USA

³Stanford Neuroscience Institute, Stanford, California, USA

⁴Symbolic Systems Program, Stanford University, Stanford, California, USA

Correspondence

Yunji Park and Vinod Menon, 401 Quarry Rd, Stanford University School of Medicine, Stanford, CA 94305, USA.

Email: yunjip@stanford.edu and menon@stanford.edu

Funding information

National Institutes of Health, Grant/Award Numbers: HD094623, HD059205, MH084164; National Science Foundation, Grant/Award Number: DRL-2024856

Abstract

Number sense is fundamental to the development of numerical problem-solving skills. In early childhood, children establish associations between non-symbolic (e.g., a set of dots) and symbolic (e.g., Arabic numerals) representations of quantity. The developmental estrangement theory proposes that the relationship between non-symbolic and symbolic representations of quantity evolves with age, with increased dissociation across development. Consistent with this theory, recent research suggests that cross-format neural representational similarity (NRS) between non-symbolic and symbolic quantities is correlated with arithmetic fluency in children but not in adolescents. However, it is not known if short-term training (STT) can induce similar changes as long-term development. In this study, children aged 7–10 years underwent a theoretically motivated 4-week number sense training. Using multivariate neural pattern analysis, we investigated whether short-term learning could modify the relation between cross-format NRS and arithmetic skills. Our results revealed a significant correlation between cross-format NRS and arithmetic fluency in distributed brain regions, including the parietal and prefrontal cortices, prior to training. However, this association was no longer observed after training, and multivariate predictive models confirmed these findings. Our findings provide evidence that intensive STT during early childhood can promote behavioral improvements and neural plasticity that resemble and recapitulate long-term neurodevelopmental changes that occur from childhood to adolescence. More generally, our study contributes to our understanding of the malleability of number sense and highlights the potential for targeted interventions to shape neurodevelopmental trajectories in early childhood.

KEYWORDS

arithmetic, individual differences, learning, multivariate neural pattern analysis, number sense, training

Research Highlights

- We tested the hypothesis that short-term number sense training induces the dissociation of symbolic numbers from non-symbolic representations of quantity in children.



- We leveraged a theoretically motivated intervention and multivariate pattern analysis to determine training-induced neurocognitive changes in the relation between number sense and arithmetic problem-solving skills.
- Neural representational similarity between non-symbolic and symbolic quantity representations was correlated with arithmetic skills before training but not after training.
- Short-term training recapitulates long-term neurodevelopmental changes associated with numerical problem-solving from childhood to adolescence.

1 | INTRODUCTION

Number sense, the ability to represent and manipulate non-symbolic (e.g., a set of dots) and symbolic (e.g., Arabic numerals) forms of quantities and understand their relations, is one of foundational cognitive abilities in our everyday life (Butterworth & Walsh, 2011; Gilmore et al., 2013; Jordan et al., 2022; Jordan et al., 2009; Rousselle & Noel, 2007; Siegler et al., 2012). Number sense is strongly predictive of higher-level math problem-solving skills and plays a critical role in mathematical skill development (Ansari, 2008; Ansari & Dhital, 2006; Bulthé et al., 2018; Schwartz et al., 2021). Here, we focus on understanding the mechanisms underlying numerical skill acquisition in response to short-term training (STT), and whether it can recapitulate learning that occurs over the course of development. Knowledge of whether an effective STT can facilitate learning and brain plasticity in a similar way as long-term learning that occurs during normative development has the potential to inform the role of learning experience in cognitive development and help develop more effective interventions.

According to the “developmental estrangement account,” symbolic quantity representations become dissociated from non-symbolic quantity representations over the course of development as individuals master the relations between non-symbolic and symbolic numbers (Lyons et al., 2012; Nakai et al., 2023; Schwartz et al., 2021). A recent study by Schwartz et al. investigated this hypothesis in a neurodevelopmental context. The researchers found that the degree of *cross-format* neural representational similarity (cross-format NRS) between non-symbolic and symbolic quantity is correlated with arithmetic skills in children but not in adolescents. In children, significant associations between cross-format NRS and arithmetic skills were found in multiple brain regions, including frontal–parietal cortical regions consistently implicated in numerical information processing (Arsalidou et al., 2018; Sokolowski et al., 2017). These findings suggest developmental changes in the association between the neural measure of number sense and arithmetic skills. However, it remains an open question whether a short-term learning focused on enhancing individuals’ understanding of quantity (number sense) can induce changes in the relation between neural mechanisms of number sense and arithmetic skills.

In the present study, we used a theoretically motivated 4-week number sense intervention designed to enhance mappings between

non-symbolic and symbolic numbers, combined with a multivariate neural pattern analysis approach. Leveraging this neural pattern analysis approach, our goal was to determine whether short-term training can induce neural changes that resemble neurodevelopmental changes observed from childhood to adolescence. Building upon findings from a previous long-term developmental (LTD) study (Schwartz et al., 2021), which included 53 children and 48 adolescents, we sought to recapitulate and extend these results within a STT context (Figure 1a). We reasoned that even in the absence of LTD changes and learning experiences, an effective STT intervention could induce comparable patterns of brain and behavioral changes to those observed over the course of long-term development. Specifically, we investigated whether findings from before and after training in a STT sample, which included 40 children aged 7–10 years, could recapitulate findings from children and adolescents in the LTD study, respectively. To establish a meaningful basis for evaluating comparable findings between the two samples at each respective time point or developmental stage, we employed the same tasks in the STT sample as those used in the LTD study. To investigate potential changes in underlying representations of mapping between two number formats induced by our intervention, we examined cross-format NRS, indexing *neural relations* between non-symbolic and symbolic representation of quantity, similar to the approach in the LTD study (Schwartz et al., 2021).

In the STT study, each participant completed two fMRI sessions, one session before training and another after training, during which they performed number comparison tasks in both non-symbolic and symbolic formats. Additionally, participants underwent a standardized test of arithmetic fluency before and after training. We first investigated whether the short-term number sense training could lead to improvements in behavioral performance on number comparison tasks in both non-symbolic and symbolic formats. Previous intervention studies targeting children’s understanding of quantity by mapping symbolic numbers to non-symbolic quantities have reported significant performance gains (Honore & Noel, 2016; Michels et al., 2018; Obersteiner et al., 2013; Wilson et al., 2006). We hypothesized that children in our STT sample would exhibit enhanced performance across both non-symbolic and symbolic number comparison tasks as a result of the intervention.

Next, we examined whether the short-term number sense training could influence the relation between performance on non-symbolic

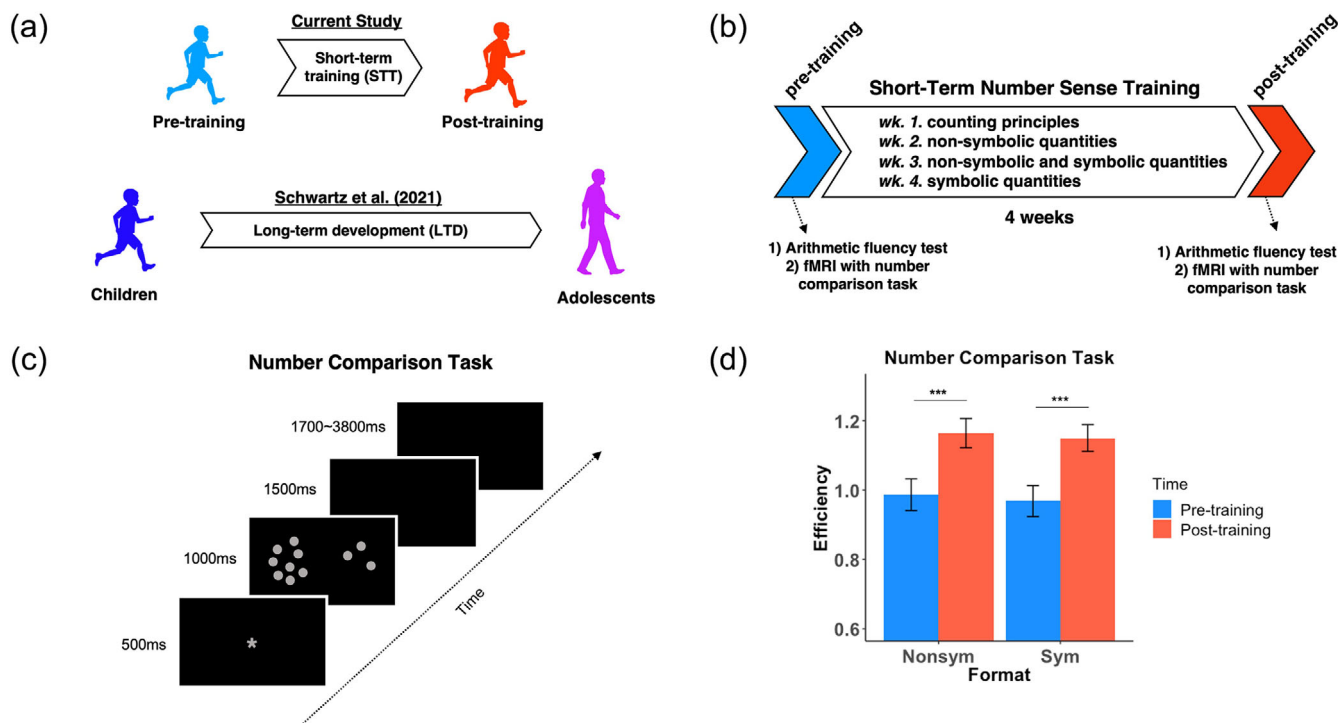


FIGURE 1 Short-term training (STT) and long-term developmental (LTD) study overview and behavioral results from STT. (A) STT and LTD samples. In the current study, we compared training-induced learning in a 4-week training study (STT) to neurodevelopmental differences between children and adolescents in Schwartz et al. (2021) (LTD). (B) STT study design. First, children's arithmetic fluency was assessed from Math Fluency subtest from Woodcock–Johnson–III (WJ-III; Woodcock et al., 2001). On a separate day, children underwent an fMRI scan session during which they had to determine which of two quantities (presented in dot arrays [or Arabic numerals] for non-symbolic [or symbolic] condition) is larger. Upon the completion of fMRI scan, children went through intensive 4 weeks of one-on-one number sense training with a tutor, focusing on improving mapping between non-symbolic and symbolic numerical quantity representations. The training occurred three times per week with each session taking ~60 min. After training, children underwent a second fMRI scan session and completed a second WJ-III Math Fluency subtest. (C) fMRI tasks. In the fMRI session, participants performed non-symbolic and symbolic number comparison tasks in separate runs. The figure depicts an example trial of a non-symbolic comparison task. Participants responded which side had the larger quantity after the onset of the presentation of the pair of quantities and before the end of the inter-trial interval. Neural representational similarity (NRS) between brain response patterns of the numerical distance effect (near–far distance) in non-symbolic and symbolic number comparison tasks was computed (see details in Methods). (D) Behavioral performance improvements in response to STT. Increased performance efficiency was observed in number comparison task in both non-symbolic (Nonsym) and symbolic (Sym) formats. Performance efficiency was measured by dividing accuracy by median reaction time, with higher scores indicating higher efficiency. $p^{***} < 0.001$.

and symbolic number comparison tasks. Previous work observed that children, not adolescents, exhibited a significant association in performance between these tasks (Schwartz et al., 2021), which suggests a developmental shift from mapping to dissociation in task performance across number formats, in line with the developmental estrangement account. We hypothesized that training-induced changes in the STT sample would resemble and recapitulate developmental differences observed between children and adolescents in the LTD sample. Specifically, we predicted that children in our STT sample would demonstrate a significant correlation between non-symbolic and symbolic number comparison task performance before training but not after training.

We then sought to determine whether the neurodevelopmental trajectories of the relation between cross-format NRS and arithmetic skills, as observed in the LTD study (Schwartz et al., 2021), could be recapitulated following a STT intervention. We hypothesized that our 4-week number sense training would facilitate learning-related neural

plasticity in the STT sample, resembling the neural plasticity observed across development in the LTD sample. Specifically, we investigated two key aspects regarding the relation between cross-format NRS and arithmetic fluency. First, we examined whether the relation between cross-format NRS and arithmetic fluency in the STT sample prior to training was comparable to that observed in children from the LTD sample. Second, we examined whether the relation between cross-format NRS and arithmetic fluency in the STT sample post-training resembled that found in adolescents from the LTD sample. By evaluating these criteria, we aimed to determine if findings from the STT sample recapitulated those observed in the LTD sample.

Our findings elucidate the effects of STT on children's number sense development, advance our understanding of shared neural representations between symbolic and non-symbolic quantity and their links with arithmetic fluency, and clarify whether STT could recapitulate long-term neurodevelopmental changes observed from

TABLE 1 List of comparison stimuli in the current study.

Condition (relative difficulty)	Big near (1)	Little near (2)	Big far (3)	Little far (4)
Stimuli repeated 4x in each task (ratio):	6:7 (0.86)	1:2 (0.50)	3:8 (0.38)	1:6 (0.17)
	8:9 (0.89)	3:4 (0.75)	4:9 (0.44)	2:7 (0.29)
	7:6 (0.86)	2:1 (0.50)	8:3 (0.38)	6:1 (0.17)
	9:8 (0.89)	4:3 (0.75)	9:4 (0.44)	7:2 (0.29)

Note: Size of stimuli (big or little) was determined by the sum of quantities in each pair.

childhood to adolescence. Our study offers valuable insights into the plasticity of neural representations induced by training and their downstream impact on children's problem-solving abilities.

2 | MATERIAL AND METHODS

2.1 | Participants

STT sample: Training group. A total of 66 second or third grade children recruited from multiple school districts participated in the training study as a part of a larger study including children with and without mathematical difficulties (Karraker et al., 2018). Among these participants, a total of 24 participants were excluded from data analysis due to excessive head movement ($n = 11$), poor image quality ($n = 3$), poor co-registration of brain images to a template ($n = 1$), brain abnormality ($n = 1$), low birth weight ($n = 1$), missing data ($n = 5$), or task administration error ($n = 2$). Two additional participants were excluded during the process of matching general cognitive abilities between children with and without mathematical difficulties in the larger study where current sample was obtained. Therefore, the final sample consisted of 40 children (21 females, $Mean_{age} = 8.18$, $SD_{age} = 0.58$).

STT sample: No-contact control group. As part of a larger study, a total of 30 second or third grade children were recruited to participate as no-contact controls for the training study. These participants completed all aspects of the training study except for the training protocol. Among 30 participants, a total of 6 participants were excluded due to missing behavioral data ($n = 3$) or below chance level performance in either non-symbolic or symbolic number comparison tasks at pre- or post-training session ($n = 3$). Therefore, the final sample for behavioral data analysis for no-contact control group included 24 children (15 female, $Mean_{age} = 8.29$, $SD_{age} = 0.60$). No additional neuroimaging data analysis was performed for no-contact controls due to insufficient sample with high-quality fMRI data across time points.

Additional details on demographics of participants in the STT sample are included in [Supplementary Methods](#). Informed written consent was obtained from the child's legal guardian and the protocol was approved by the Institutional Review Board.

LTD sample. We used the previous dataset of Schwartz et al. (2021). The final sample of Schwartz et al. consisted of 53 children (28 females, $Mean_{age} = 8.2$, $SD_{age} = 0.65$) and 48 adolescents (26 females, $Mean_{age} = 18.2$, $SD_{age} = 1.56$). Additional details are described in Schwartz et al. (2021).

2.2 | Experimental procedures

The STT study aimed to examine cognitive and brain plasticity following number sense training. Children completed cognitive assessments, which included an arithmetic fluency test, and brain imaging scans, which included number comparison fMRI tasks, before and after a 4-week number sense training (Figure 1b,c). Children and adolescents in the LTD study (Schwartz et al., 2021) completed identical cognitive and brain measures.

2.2.1 | Non-symbolic and symbolic number comparison fMRI tasks

Children completed the non-symbolic and symbolic number comparison tasks in the MRI scanner wherein they determined the larger between two non-symbolic or symbolic numbers (Figure 1c). We used a 2×2 experimental design accounting for both the size of the number pair (little, big) and the distance between the numbers of the pair (near, far), resulting in 16 trials per condition (Table 1). In half of the trials, the sum of the pair was >10 ("big" comparison), and in the other half, it was <10 ("little" comparison). In half of the trials, the distance between the numbers was 1 ("near" distance), and in the other half, the distance was 5 ("far" distance). Quantities between 1 and 9, excluding 5, were used.

On each trial, a fixation appeared for 500 ms followed by a pair of quantities, which remained visible for 1000 ms before disappearing, leaving a blank screen for 1500 ms to fill up the response phase. Stimuli were presented simultaneously on each side of the screen and consisted of a pair of dot arrays in the non-symbolic task and a pair of Arabic numbers in the symbolic task. The duration for stimulus presentation was relatively short in order to avoid counting in the non-symbolic condition. For the non-symbolic number comparison task, we controlled (i) the total area covered by each array of dots for the half of the trials and (ii) the average size of dots in each array for the other half of the trials to control for potential confounds with the number of items. The size of dots in the non-symbolic task varied, with dots being spaced with similar density without apparent pattern. Each stimulus pair was presented four times, once with the larger number on the left, and once with the larger number on the right.

With a button box (2-button fiber optic response pad, Curdes system: www.Curdes.com), children indicated their judgment of which



side of the screen had larger quantity via button press—the left button if the larger number was on the left side or the right button if the larger number was on the right side. Stimuli were presented using E-Prime 2.0 (Schneider et al., 2002) and displayed using an LCD projector and a back-projection screen in the scanner suite. The intertrial interval between trials was jittered randomly between 1.7 and 3.8 seconds. A total of 64 trials were presented in each run, which took about 6 minutes. The order of trials and runs (symbolic or non-symbolic) was randomized across participants.

2.2.2 | Arithmetic fluency

Arithmetic problem-solving skills were assessed using standardized scores from the Math Fluency subtest of the Woodcock Johnson-III (Woodcock et al., 2001). This subtest is a timed (3 min.) pen-and-paper test that measures individuals' ability to accurately solve simple addition, subtraction, and multiplication as quickly as possible. All problems were presented vertically and involved operands from 0 to 10. We then utilized standardized scores based on population norms provided by the Woodcock Johnson-III for our analysis.

2.2.3 | Number sense training

In the STT study, children participated in a 4-week number sense training, which was designed to enhance their fundamental understanding of quantity in non-symbolic and symbolic forms and the relations between non-symbolic and symbolic quantities. Our number sense training primarily focused on enhancing children's understanding of cross-format mapping to develop proficiency with exact symbolic numbers (Obersteiner et al., 2013; Tobia et al., 2021), rather than targeting refinement of approximate non-symbolic representation of large quantities (Lau et al., 2021; Suárez-Pellicioni & Booth, 2018). Specifically, children's training activities progressed from week to week: children learned and practiced basic counting principle in week 1, non-symbolic quantity comparison in week 2, non-symbolic and symbolic quantity comparison in week 3, and symbolic quantity comparison in week 4. In each training session, a trained tutor used various instruction methods, including the use of physical manipulatives, computer games, and review worksheets. Training occurred three times a week over the course of 4 weeks. Training activities from each week are described below.

Week 1. The week 1's training sessions focused on the counting principle. The child was introduced to counting principles with some video clips demonstrating accurate or inaccurate counting using erasers and sock puppets. After the lesson, children played *Restaurant Game* (Blair, 2013a, 2013b), in which they were asked to count the number of dishes to cook for animals presented in the screen. In the last session of week 1, children were asked to complete a 42-problem review worksheet, in which they were asked to count the number of animals and to draw a circle around, match, or write the number they counted.

Week 2. The week 2's sessions focused on the comparison of non-symbolic numbers using sets of erasers. Children played *Math Circles* wherein they determined which of two *Math Circles* presented on the table had more erasers. This lesson aimed to get children familiarized with the concept of number comparison using non-symbolic quantities. Starting from week 2, children played a computerized game, an adapted version of *Number Race* (Wilson et al., 2006). The contents of each week's game were corresponded to the progression of training activities and did not include arithmetic training. For example, in week 2, the game mainly focused on comparisons of non-symbolic quantities. Children also played two interactive games with the tutor: (1) *Math War* (Iuculano et al., 2015) wherein children compared two quantities and (2) *Comparing speed*, in which children identified the quantity of one value above or below the given quantity. In the *Math War* game, depending on the week's focus (non-symbolic and/or symbolic numbers), a deck of card with non-symbolic and/or symbolic number format was each assigned to the child and the tutor, both of whom chose a card from their deck one at a time. The child was asked to write down the number on his/her own card and the one on the tutor's card, and then asked to determine the card with a larger quantity. In the *Comparing speed* game, the child and tutor each took five cards from their own deck of cards once after the tutor laid four cards with quantities from 4 to 7 on the table. If the quantity on the card each took was below or above the quantity on the table, they placed the card on the top of the card on the table. The game finished when the child placed all the cards from their own deck. As *Math War*, the format of quantity presented on the card varied depending on the week's focus. In each session in week 2, children completed a review worksheet, which consisted of 24 non-symbolic number comparison problems.

Week 3. The week 3's sessions had similar paradigm as week 2's sessions, but with a focus on integrating between symbolic numbers and non-symbolic quantities. In *Math Circles*, children compared a set of erasers in one Math Circle to a card presenting a symbolic number (Arabic numeral) in another Math Circle. Next, children played *Number Race* focusing on comparisons of non-symbolic quantities and symbolic numbers. Following *Number Race*, children also played *Math War* and *Comparing speed* with one deck of cards presenting non-symbolic numbers and the other deck presenting symbolic numbers. In each session in week 3, children completed a 48-problem review worksheet, in which they determined a larger quantity between non-symbolic and symbolic numbers.

Week 4. The week 4's sessions focused on enhancing understanding and fluent use of symbolic numbers. Children played a number ordering version of *Beat Your Score* (Chang et al., 2019) wherein the child ordered four decks of cards for three times (the tutor shuffled the card for each trial) with the aim of "beating" the time taken for the previous trial. The quantities on the cards were presented with non-symbolic, symbolic, or mixed formats. As in weeks 2 and 3, children played *Number Race* focusing on comparisons of symbolic numbers, and also played *Math War* and *Comparing speed* with the cards presenting only symbolic numbers. In each session in week 4, children completed a review worksheet, which consisted of 24 symbolic number comparison problems.

Additional information on the tutoring protocol can be found in Chang et al. (2022).

2.3 | Statistical analysis

2.3.1 | Behavioral data analysis

For symbolic and non-symbolic number comparison tasks, trials with response times lower than 150 ms were excluded from analysis of task performance, which constituted 2.42% of all trials for non-symbolic comparison (pre: 1.88%, post: 2.96%) and 1.99% of all trials for symbolic comparison (pre: 2.27%, post: 1.72%). To assess performance on non-symbolic and symbolic number tasks, we computed an efficiency score obtained by accuracy divided by median reaction times (RTs) for each task to better capture variability in performance across different dimensions of RT and accuracy, while minimizing Type I error by reducing the number of statistical tests required to test our hypotheses. To address potential concern about the limitation of efficiency score (Bruyer & Brysbaert, 2011), we additionally report findings of performance assessed by RT and accuracy (Supplementary Results). To quantify the discriminability between pairs of quantities with varying distances, we computed the numerical distance effect (NDE) by calculating the difference between near distance trials (e.g., 2 vs. 3) and far distance trials (e.g., 2 vs. 7). This approach is widely used to compute NDE effects in both non-symbolic and symbolic number comparison tasks (Ashkenazi et al., 2013; Brannon & Merritt, 2011; Holloway & Ansari, 2009; Moyer & Landauer, 1967).

ANOVA and *t*-tests were performed to examine training-induced changes in behavioral performance, assessed by efficiency scores or NDE. For efficiency scores and NDE, we performed a 2 × 2 within-subject ANOVA with factors Task (non-symbolic, symbolic) and Time (pre, post) and follow-up two-tailed paired *t*-tests to examine changes in behavioral performance in response to training or differences in performance between task formats. Planned paired *t*-test between pre- and post-training was performed to assess changes in arithmetic fluency. Pearson correlations were used to assess the relations between non-symbolic and symbolic efficiency scores or NDE at pre- and post-training and Fisher's *z*-test was employed to compare correlations between pre- and post-training. We compared the behavioral findings in our STT sample with those in the LTD sample. To report effect sizes, we used a generalized eta squared (η^2) for ANOVA (Bakeman, 2005) and Cohen's *d* for *t*-tests (Cohen, 1992).

We calculated Bayes factor (BF_{10}) to evaluate the presence or absence of evidence for the test hypothesis (Keyes et al., 2020). BF_{10} values exceeding 3 indicate evidence for the test hypothesis, while values between 0.33 and 3 suggest no clear evidence, and values lower than 0.33 indicate evidence for the null hypothesis. Additionally, higher values indicate stronger evidence for the test hypothesis. Specifically, BF_{10} values between 3 and 10 represent moderate evidence; values between 10 and 30 indicate strong evidence; values between 30 and 100 suggest robust evidence; and values exceeding

100 indicate extreme evidence (Faulkenberry et al., 2020; Jeffreys, 1961). Statistical significance was determined using both *p*-values from frequentist statistics and BF_{10} values from Bayesian statistics (i.e., $p < 0.05$ and $BF_{10} > 3$). Nonparametric equivalent ANOVA, *t*-tests, and correlations were performed to confirm the findings from parametric analyses for non-normally distributed data (see Supplementary Methods and Results).

2.3.2 | fMRI data acquisition and preprocessing

Task-based functional images were acquired on a 3T GE scanner (General Electric, Milwaukee, WI) using an eight-channel GE head coil. A T1-weighted structural image was acquired at both pre- and post-training scan sessions to facilitate registering each participant's data to standard space. Images were preprocessed and analyzed using SPM12 (Ashburner et al., 2014).

During image acquisition, head movement was minimized using additional pads and pillows around children's head. A T2*-sensitive gradient echo spiral in-out pulse sequence (Glover & Lai, 1998) was acquired with the following parameters: repetition time (TR) = 2000 ms, echo time (TE) = 30 ms, flip angle = 80°, field of view (FOV) = 220 mm, matrix size = 64 × 64, resolution = 3.44 × 3.44 × 4.5 mm³, interleaved. A total of 31 axial slices were acquired, 4 mm in thickness and 0.5 mm in spacing, covering the whole brain. A T1-weighted, 132 slice high-resolution structural image was acquired (slice thickness 1 mm; in-plane resolution: 256 × 256, voxel size = 1.5 × 0.9 × 1.1 mm³).

Among acquired images, the first five volumes were discarded to allow for signal equilibration. The preprocessing pipeline included realignment, slice-timing correction, co-registration to subjects' T1 and normalization to a standardized 2-mm adult Montreal Neurological Institute (MNI) 152 template in SPM and smoothing using a 6-mm full-width half-maximum Gaussian kernel to decrease spatial noise. We used an adult MNI template similar to a standard practice in the field of developmental cognitive neuroscience (Cantlon et al., 2006; Kersey & Cantlon, 2017; Nakai et al., 2023; Perrachione et al., 2016), considering that there is minimal functional activation differences between the brains of young children and adults (Burgund et al., 2002; Kang et al., 2003; Szaflarski et al., 2012).

Transitional (*x*, *y*, *z*) and rotational (pitch, roll, yaw) movement parameters were generated from the realignment procedure. Volumes with >0.5 voxel scan-to-scan displacement along linear or rotational axes were deweighted, as well as volumes with >5% change in global signal. The proportion of volumes with scan-to-scan displacement higher than 0.5 voxel did not exceed 10% across tasks. Movement did not exceed 12mm in any rotational and translational axes. Mean scan-to-scan displacement did not exceed 0.8mm. For participants included in data analysis (see Participants section), there were no differences between symbolic and non-symbolic tasks on movement parameters before ($ps > 0.282$) and after training ($ps > 0.247$), and no differences between the pre- and post-training scans on these parameters on symbolic ($ps > 0.073$) and non-symbolic tasks ($ps > 0.237$).



2.3.3 | fMRI data analyses

First-level statistical analysis

Following standard preprocessing of fMRI data, a general linear model (GLM) implemented in SPM12 was used to determine task-related brain responses (e.g., luculano et al., 2015).

In the first-level analysis, brain responses representing correct trials for each condition (i.e., big near, big far, little near, little far) were modeled using boxcar functions of 2500ms corresponding to the length of a trial convolved with a canonical hemodynamic response function and a temporal derivative to account for voxel-wise latency differences in hemodynamic response. An error regressor was also included in the model to account for the influence of incorrect trials. Additionally, both transitional and rotational head movement parameters (six parameters) were included as regressors of no interest. Serial correlations were accounted for by modeling the fMRI time series as a first-degree autoregressive process. The GLM was applied to non-symbolic and symbolic comparison tasks separately. Voxel-wise contrast maps were generated for each participant for each task. For both comparison tasks, the contrast of interest was the near versus far distance condition (collapsing “big” and “little” conditions in each distance) to assess neural representations of quantity in a similar way as behavioral NDE.

Multivoxel neural representational similarity (NRS) analysis

We next conducted NRS analysis based on neural distance effects to assess neural representations of quantities while carefully controlling for low-level perceptual features, motor responses, and mental activity associated with resting baseline in both non-symbolic and symbolic tasks parallel to Schwartz et al. (2021). Multivariate spatial correlation of brain activity patterns between non-symbolic and symbolic comparison tasks was computed across the whole brain for each individual at each time point (pre- and post-training). Using a searchlight mapping method (Kriegeskorte et al., 2006), we obtained cross-format NRS of the NDE in the neighborhood surrounding each voxel of each individual's brain. Specifically, a 6-mm spherical region centered on each voxel was selected, and then cross-format similarity was computed within the sphere using Pearson correlation between voxel-wise brain activation (beta-weights). Searchlight maps were then created for every individual by going through every voxel across the whole brain and assigned NRS values in the center voxel of 6-mm spherical regions.

We focused on the regions of interest (ROIs) that showed significant relation between cross-format NRS and arithmetic fluency in children in the LTD cohort of Schwartz et al. (2021). This included bilateral IPS, precentral gyrus (PreCG), hippocampus (HIPP), thalamus, left Premotor gyrus (preMotor), left insula, left middle frontal gyrus (MFG), and right central operculum (COper) (Tables 2–6 present peak coordinate for each region). Each ROI was generated as a spherical region with 6-mm radius centered on each peak coordinate, which corresponded to the size of ROIs in the searchlight NRS analysis and contained approximately 17 voxels. We computed Pearson correlations between cross-format NRS in these regions and arithmetic fluency at each time

point (pre- and post-training). Correlations between cross-format NRS and arithmetic fluency in our STT sample were compared with the LTD sample with Fisher's z-test. Similar to behavioral data analysis, both *p*-values from frequentist statistics and BF_{10} values were used to determine the significance of statistical tests. To examine multivariate brain-behavioral relations, we performed a multilinear regression predicting arithmetic fluency scores. Cross-format NRS values from a total of 15 ROIs were z-transformed and entered into the model as predictors (see [Supplementary Results](#) for additional confirmatory whole brain analyses).

3 | RESULTS

3.1 | Number sense training improves children's non-symbolic and symbolic number comparison performance efficiency

We first examined whether our 4-week training improved number sense in both non-symbolic and symbolic formats in the training group in the STT sample. Efficiency score (accuracy/median RT) and NDE (near–far efficiency) were used to assess performance efficiency and discriminability of quantity, respectively (see details in *Methods*). We conducted a 2 × 2 within-subject ANOVA with factors Format (non-symbolic, symbolic) and Time (pre, post) for both measures.

Efficiency score. We found a significant main effect of Time ($F(1,39) = 30.397, p < 0.001, \eta^2 = 0.101, BF_{10} = 9.0e7$) but no significant main effect of Format ($p = 0.501, BF_{10} = 0.20$) or Format × Time interaction ($p = 0.903, BF_{10} = 0.225$) on efficiency score for number comparison task. Follow-up paired *t*-test comparing pre- and post-training confirmed significant increases in efficiency scores in both non-symbolic ($t(39) = 4.216, p < 0.001, \text{Cohen's } d = 0.639, BF_{10} = 176.01$) and symbolic ($t(39) = 5.902, p < 0.001, \text{Cohen's } d = 0.686, BF_{10} = 2.3e4$) number comparison tasks (Figure 1d).

We then conducted similar analyses using accuracy and RT measures (see [Supplementary Results](#) for details). We found a significant main effect of Format on accuracy and significant main effect of Time and Format and Format × Time interaction for RTs ($ps < 0.05$). No other main effects or interactions were significant for accuracy ($ps > 0.061$). Follow-up *t*-tests revealed higher accuracy for symbolic compared to non-symbolic number comparison and significant decreases in RTs after training for both non-symbolic and symbolic number comparisons ($ps < 0.001$). These findings suggest that training-induced changes in RTs may have primarily contributed to significant changes in performance efficiency.

Additional analysis with the LTD sample demonstrated significantly higher non-symbolic and symbolic number comparison task efficiency in adolescents compared to children ([Supplementary Results](#)), which indicates that number comparison efficiency in both formats may improve over long-term development and learning.

NDE. We found a significant main effect of Format ($F(1,39) = 93.394, p < 0.001, \eta^2 = 0.346, BF_{10} = 3.7e16$) but no significant main effect of

TABLE 2 Regional correlations of neural representational similarity (NRS) between non-symbolic and symbolic quantity and arithmetic fluency scores in children at pre- and post-training in the short-term training (STT) sample.

ROI [MNI coordinates]:	Pre-training			Post-training			Difference in correlation		
	<i>r</i>	<i>p</i>	<i>BF</i> ₁₀	<i>r</i>	<i>p</i>	<i>BF</i> ₁₀	<i>z</i>	<i>p</i>	<i>BF</i> ₁₀
L IPS/LOC [−22 −62 48]	0.48**	0.008	29.19	0.02	0.585	0.35	2.17*	0.030	1.58
L IPS [−36 −48 48]	0.43**	0.005	10.57	0.10	0.552	0.41	1.49	0.136	0.59
L IPS [−26 −52 48]	0.41**	0.002	7.48	−0.09	0.906	0.40	2.27*	0.023	0.80
R IPS [30 −52 38]	0.39**	0.012	5.64	0.29	0.075	1.43	0.51	0.607	0.26
R IPS [42 −44 46]	0.38	0.016	4.47	0.23	0.154	0.86	0.79	0.430	0.29
L PreCG [−2 −24 70]	0.47**	0.002	21.38	0.003	0.986	0.35	2.16*	0.031	1.22
R PreCG [4 −22 62]	0.37*	0.020	3.83	0.001	0.990	0.35	1.64	0.102	0.67
L HIPP [−20 −16 −14]	0.35*	0.026	3.15	0.14	0.377	0.50	1.07	0.285	0.36
R HIPP [24 −20 −16]	0.34*	0.034	2.55	0.20	0.221	0.68	0.58	0.560	0.31
L Thalamus [−16 −18 0]	0.44**	0.004	12.56	0.24	0.137	0.93	1.06	0.288	0.29
R Thalamus [14 −14 2]	0.49**	0.001	36.21	0.22	0.170	0.81	1.31	0.191	0.31
L preMotor [−16 −6 62]	0.45**	0.004	13.53	−0.09	0.564	0.41	2.21*	0.027	2.12
L Insula [−38 −8 −10]	0.54***	<0.001	109.81	−0.14	0.406	0.48	3.04**	0.002	7.00
L MFG [−34 2 64]	0.56***	<0.001	162.16	0.32*	0.043	2.14	1.27	0.203	0.24
R COper [42 −8 16]	0.52**	0.001	67.66	−0.08	0.626	0.39	2.71**	0.007	4.09

Notes: Regions of interest (ROIs) were defined from Schwartz et al. (2021). *BF*₁₀ values greater than 3 (bolded) provide evidence for the existence of correlation, while *BF*₁₀ values between 0.33 and 3 provide absence of evidence and *BF*₁₀ values lower than 0.33 provide evidence of no correlation. *BF*₁₀ values between 3 and 10 represent moderate evidence; values between 10 and 30 indicate strong evidence; values between 30 and 100 suggest robust evidence; and values exceeding 100 indicate extreme evidence. All *z* values indicate effect sizes of difference in correlation between pre- and post-training.

Abbreviations: COper, central operculum; HIPP, hippocampus; IPS, intraparietal sulcus; L, left; LOC, lateral occipital cortex; MFG, middle frontal gyrus; PreCG, precentral gyrus; preMotor, premotor cortex; R, right.

p* < 0.05, *p* < 0.01, ****p* < 0.001.

Time (*p* = 0.148, *BF*₁₀ = 0.32) or Format × Time interaction (*p* = 0.165, *BF*₁₀ = 0.42). Follow-up two sample *t*-test comparing formats confirmed significant larger distance effects in non-symbolic comparison at both pre- (*t*(62.84) = −5.460, *p* < 0.001, Cohen's *d* = −1.22, *BF*₁₀ = 2.4e4) and post-training (*t*(77.89) = −7.513, *p* < 0.001, Cohen's *d* = −1.68, *BF*₁₀ = 8.7e7). Additional analysis with the LTD sample demonstrated similar patterns of non-significant developmental differences in symbolic NDE (though this was not the case for non-symbolic NDE; see details in [Supplementary Results](#)).

These results show that our number sense training leads to similar levels of performance improvements across formats.

3.2 | Number sense training induces dissociation between non-symbolic and symbolic quantity discrimination ability

We next examined correlations between non-symbolic and symbolic task performance in the training group in the STT sample to address whether training alters the extent to which efficiency score or NDE relates between number formats.

Efficiency score. We observed that the efficiency score is significantly correlated between formats before training (*r* = 0.827, *p* < 0.001, *BF*₁₀ = 5.2e7) and also after training (*r* = 0.694, *p* < 0.001,



TABLE 3 Regional correlations of neural representational similarity (NRS) between non-symbolic and symbolic quantity and arithmetic fluency scores in children at pre-training in the short-term training (STT) sample and children in the long-term developmental (LTD) sample.

ROI [MNI coordinates]:	Pre-training (STT sample)			Children (LTD sample)			Difference in correlation		
	<i>r</i>	<i>p</i>	<i>BF</i> ₁₀	<i>r</i>	<i>p</i>	<i>BF</i> ₁₀	<i>z</i>	<i>p</i>	<i>BF</i> ₁₀
L IPS/LOC [−22 −62 48]	0.48**	0.008	29.19	0.39**	0.006	12.21	0.57	0.570	0.23
L IPS [−36 −48 48]	0.43**	0.005	10.57	0.36**	0.009	7.14	0.40	0.686	0.27
L IPS [−26 −52 48]	0.41**	0.002	7.48	0.38**	0.004	10.14	0.19	0.848	0.23
R IPS [30 −52 38]	0.39*	0.012	5.64	0.42**	0.002	29.21	0.17	0.863	0.25
R IPS [42 −44 46]	0.38*	0.016	4.47	0.35*	0.010	6.27	0.14	0.886	0.25
L PreCG [−2 −24 70]	0.47**	0.002	21.38	0.39**	0.004	13.10	0.46	0.647	0.23
R PreCG [4 −22 62]	0.37*	0.020	3.83	0.38**	0.005	11.28	0.08	0.939	0.26
L HIPP [−20 −16 −14]	0.35*	0.026	3.15	0.42**	0.002	25.76	0.36	0.718	0.27
R HIPP [24 −20 −16]	0.34*	0.034	2.55	0.35**	0.009	6.71	0.10	0.923	0.27
L Thalamus [−16 −18 0]	0.44**	0.004	12.56	0.46**	0.001	63.27	0.08	0.935	0.23
R Thalamus [14 −14 2]	0.49**	0.001	36.21	0.54***	<0.001	914.36	0.30	0.761	0.23
L preMotor [−16 −6 62]	0.45**	0.004	13.53	0.43**	0.001	35.98	0.07	0.944	0.22
L Insula [−38 −8 −10]	0.54***	<0.001	109.81	0.47***	<0.001	98.40	0.44	0.660	0.19
L MFG [−34 2 64]	0.56***	<0.001	162.16	0.42**	0.002	24.41	0.86	0.392	0.23
R COper [42 −8 16]	0.52**	0.001	67.66	0.45**	0.001	59.57	0.42	0.673	0.21

Note: Regions of interest (ROIs) were defined from Schwartz et al. (2021). *BF*₁₀ values greater than 3 (bolded) provide evidence for the existence of correlation, while *BF*₁₀ values between 0.33 and 3 provide absence of evidence and *BF*₁₀ values lower than 0.33 provide evidence of no correlation. *BF*₁₀ values between 3 and 10 represent moderate evidence; values between 10 and 30 indicate strong evidence; values between 30 and 100 suggest robust evidence; and values exceeding 100 indicate extreme evidence. All *z* values indicate effect sizes of difference in correlation between the STT and LTD samples.

Abbreviations: COper, central operculum; HIPP, hippocampus; IPS, intraparietal sulcus; L, left; LOC, lateral occipital cortex; MFG, middle frontal gyrus; PreCG, precentral gyrus; preMotor, premotor cortex; R, right.

p* < 0.05, *p* < 0.01, ****p* < 0.001.

*BF*₁₀ = 1.7e4), which indicates association between the two number formats remained after training. Next, we compared correlation between efficiency score across formats between the STT and LTD samples. We found that correlation between non-symbolic and symbolic efficiency at pre-training in STT sample was not significantly different from correlation in children in the LTD sample (*z* = 0.74, *p* = 0.460) and that the correlation at post-training in the STT sample was not significantly different from correlation in adolescence in the LTD sample (*z* = 0.96, *p* = 0.340). Although these results did not support training-induced dissociation of symbolic number format from non-symbolic formats in the STT sample, the results indicate that the degree of cross-format correlation in efficiency of the STT sample at

pre- and post-training resembled those of children and adolescents from the LTD sample, respectively.

NDE. We observed that the NDE is significantly correlated between formats before training (*r* = 0.424, *p* = 0.006, *BF*₁₀ = 9.34), but not after training (*r* = 0.261, *p* = 0.103, *BF*₁₀ = 1.14; Supplementary Table 1), which indicates dissociation between the two number formats. Next, we compared correlation between NDEs across formats between the STT and LTD samples. We found that correlation between non-symbolic and symbolic NDE at pre-training in the STT sample was not significantly different from correlation in children in the LTD sample (*z* = 0.26, *p* = 0.79, *BF*₁₀ = 0.23). Furthermore, correlation at post-training in the STT sample was not significantly different from



TABLE 4 Regional correlations of neural representational similarity (NRS) between non-symbolic and symbolic quantity and arithmetic fluency scores in children at post-training in the short-term training (STT) sample and adolescents in the long-term developmental (LTD) sample.

ROI [MNI coordinates]:	Post-training (STT sample)			Adolescents (LTD sample)			Difference in correlation		
	<i>r</i>	<i>p</i>	<i>BF</i> ₁₀	<i>r</i>	<i>p</i>	<i>BF</i> ₁₀	<i>z</i>	<i>p</i>	<i>BF</i> ₁₀
L IPS/LOC [−22 −62 48]	0.02	0.585	0.35	−0.12	0.629	0.44	0.30	0.761	0.48
L IPS [−36 −48 48]	0.10	0.552	0.41	0.10	0.525	0.40	0.15	0.883	0.32
L IPS [−26 −52 48]	−0.09	0.906	0.40	0.07	0.432	0.37	0.90	0.370	0.36
R IPS [30 −52 38]	0.29	0.075	1.43	0.26	0.084	1.27	0.18	0.860	0.29
R IPS [42 −44 46]	0.23	0.154	0.86	−0.05	0.744	0.35	1.41	0.158	0.76
L PreCG [−2 −24 70]	0.003	0.986	0.35	−0.10	0.531	0.40	0.61	0.544	0.41
R PreCG [4 −22 62]	0.001	0.990	0.35	−0.04	0.783	0.35	0.31	0.759	0.34
L HIPP [−20 −16 −14]	0.14	0.377	0.50	0.07	0.666	0.37	0.40	0.688	0.34
R HIPP [24 −20 −16]	0.20	0.221	0.68	0.08	0.622	0.38	0.51	0.608	0.35
L Thalamus [−16 −18 0]	0.24	0.137	0.93	−0.05	0.747	0.35	1.34	0.181	0.62
R Thalamus [14 −14 2]	0.22	0.170	0.81	0.01	0.943	0.34	0.71	0.475	0.35
L preMotor [−16 −6 62]	−0.09	0.564	0.41	−0.07	0.647	0.37	0.09	0.932	0.33
L Insula [−38 −8 −10]	−0.14	0.406	0.48	−0.06	0.690	0.36	0.44	0.661	0.38
L MFG [−34 2 64]	0.32*	0.043	2.14	<0.001	0.979	0.34	1.49	0.136	0.69
R COper [42 −8 16]	−0.08	0.626	0.39	−0.14	0.349	0.50	0.10	0.917	0.32

Note: Regions of interest (ROIs) were defined from Schwartz et al. (2021). *BF*₁₀ values greater than 3 (bolded) provide evidence for the existence of correlation, while *BF*₁₀ values between 0.33 and 3 provide absence of evidence and *BF*₁₀ values lower than 0.33 provide evidence of no correlation. *BF*₁₀ values between 3 and 10 represent moderate evidence; values between 10 and 30 indicate strong evidence; values between 30 and 100 suggest robust evidence; and values exceeding 100 indicate extreme evidence. All *z* values indicate effect sizes of difference in correlation between the STT and LTD samples.

Abbreviations: COper, central operculum; HIPP, hippocampus; IPS, intraparietal sulcus; L, left; LOC, lateral occipital cortex; MFG, middle frontal gyrus; PreCG, precentral gyrus; preMotor, premotor cortex; R, right.

**p* < 0.05.

correlation in adolescents in the LTD sample ($z = 0.47$, $p = 0.64$, $BF_{10} = 0.41$). These findings suggest that short-term number sense training led to dissociation of symbolic numbers from non-symbolic representations of quantity in a similar way as LTD changes in quantity representation.

3.3 | No significant changes in non-symbolic and symbolic number comparison task performance following 4 weeks without intervention

To examine potential changes in performance across four weeks without intervention on non-symbolic and symbolic number comparison

tasks, we performed ANOVA and follow-up *t*-tests on efficiency score and NDE in the control group for the STT sample.

Efficiency score. Using a 2×2 within-subject ANOVA with factors Format (non-symbolic, symbolic) and Time (Pre, Post), we found no significant main effect of Time ($p = 0.208$) or Format by Time interaction ($p = 0.704$). The main effect of Format was not significant for parametric analysis ($F(1, 23) = 5.104$, $p = 0.034$, $\eta^2 = 0.014$, $BF_{10} = 1.267$). These results indicate no significant gains in efficiency score of number comparison tasks across time in the control group.

NDE. Using a 2×2 within-subject ANOVA with factors Format and Time, we found a significant main effect of Format ($F(1,$



TABLE 5 Multilinear regression analysis of neural representational similarity (NRS) between non-symbolic and symbolic quantity of multiple regions predicting arithmetic fluency scores in children at pre- and post-training in the short-term training (STT) sample.

ROI [MNI coordinates]:	Pre-training				Post-training			
	Adjusted $R^2 = 0.486^{**}$, $p = 0.003$				Adjusted $R^2 = -0.039$, $p = 0.572$			
	β	t	p	vif	β	t	p	vif
L IPS/LOC [-22 -62 48]	0.26	1.07	0.296	4.65	0.14	0.58	0.568	2.22
L IPS [-36 -48 48]	0.19	1.26	0.221	1.70	0.17	0.73	0.474	2.13
L IPS [-26 -52 48]	-0.25	-1.18	0.252	3.33	-0.26	-0.88	0.390	3.24
R IPS [30 -52 38]	0.42*	2.14	0.042	2.86	0.35	1.50	0.148	2.10
R IPS [42 -44 46]	-0.13	-0.75	0.464	2.46	-0.08	-0.32	0.753	2.63
L PreCG [-2 -24 70]	0.21	1.11	0.278	2.63	-0.22	-0.81	0.427	2.85
R PreCG [4 -22 62]	0.04	0.18	0.857	4.08	0.27	0.82	0.420	4.17
L HIPP [-20 -16 -14]	0.23	1.24	0.226	2.68	-0.10	-0.51	0.617	1.45
R HIPP [24 -20 -16]	0.17	1.06	0.298	2.04	-0.07	-0.29	0.778	2.50
L Thalamus [-16 -18 0]	-0.27	-1.17	0.254	3.91	0.21	0.85	0.403	2.34
R Thalamus [14 -14 2]	0.13	0.68	0.503	2.86	0.10	0.43	0.668	2.09
L preMotor [-16 -6 62]	-0.29	-1.33	0.197	3.66	-0.21	-1.04	0.310	1.58
L Insula [-38 -8 -10]	0.36*	2.32	0.029	1.81	-0.18	-0.85	0.403	1.74
L MFG [-34 2 64]	0.12	0.68	0.504	2.38	0.46*	2.23	0.036	1.57
R COper [42 -8 16]	0.25	1.67	0.109	1.64	-0.01	-0.04	0.967	1.74

Note: Regions of interest (ROIs) were defined from Schwartz et al. (2021). All β values indicate standardized weight for each ROI.

Abbreviations: COper, central operculum; HIPP, hippocampus; IPS, intraparietal sulcus; L, left; LOC, lateral occipital cortex; MFG, middle frontal gyrus; PreCG, precentral gyrus; preMotor, premotor cortex; R, right.

* $p < 0.05$, ** $p < 0.01$.

26) = 22.291, $p = 0.0001$, $\eta^2 = 0.213$, $BF_{10} = 2.5e4$) and no significant main effect of Time ($p = 0.524$) or Format by Time interaction ($p = 0.431$). Follow-up paired t -test comparing formats confirmed significantly larger NDE in non-symbolic compared to symbolic format at both pre-training ($t(45.76) = -3.090$, $p = 0.004$, Cohen's $d = -0.892$, $BF_{10} = 11.469$) and post-training ($t(50.06) = -3.947$, $p < 0.001$, Cohen's $d = -1.14$, $BF_{10} = 93.938$), similar to observed format effect in the training group in the STT sample. These results indicate no significant gains in NDE of number comparison tasks across time in the control group.

Together, these findings suggest that 4 weeks of no intervention did not significantly change the efficiency or NDE of number comparison tasks.

3.4 | No significant changes in correlation between non-symbolic and symbolic number comparison task performance following 4 weeks without intervention

To examine potential correlations between non-symbolic and symbolic task performance across 4 weeks without intervention, we performed Pearson's correlations on efficiency score and NDE in the control group for the STT sample. Here, we observed that the efficiency score is significantly correlated between formats at both time points (pre: $r = 0.755$, $p < 0.001$, $BF_{10} = 6.5e2$, post: $r = 0.671$, $p = 0.003$, $BF_{10} = 78.92$), with no significant difference between correlations at each time point ($p = 0.580$). For NDE, we found no significant correlation between two number formats at both pre- ($r = 0.005$, $p = 0.982$,



TABLE 6 Multilinear regression analysis of neural representational similarity (NRS) between non-symbolic and symbolic formats of multiple regions predicting arithmetic fluency scores in children and adolescents in the long-term developmental (LTD) sample.

ROI [MNI coordinates]:	Children				Adolescents			
	β	<i>t</i>	<i>p</i>	<i>vif</i>	β	<i>t</i>	<i>p</i>	<i>vif</i>
L IPS/LOC [−22 −62 48]	0.05	0.28	0.783	2.74	−0.22	−0.93	0.361	2.11
L IPS [−36 −48 48]	0.19	1.36	0.183	1.70	0.06	0.24	0.812	2.14
L IPS [−26 −52 48]	−0.21	−1.14	0.261	3.11	0.18	0.59	0.559	3.43
R IPS [30 −52 38]	0.33*	2.19	0.035	2.03	0.43	1.83	0.078	1.95
R IPS [42 −44 46]	0.06	0.47	0.640	1.53	−0.31	−1.22	0.234	2.41
L PreCG [−2 −24 70]	0.04	0.25	0.803	2.51	−0.13	−0.54	0.595	2.01
R PreCG [4 −22 62]	0.10	0.49	0.624	3.84	0.17	0.62	0.543	2.73
L HIPP [−20 −16 −14]	0.21	1.50	0.142	1.83	0.17	0.62	0.537	2.69
R HIPP [24 −20 −16]	0.13	0.90	0.373	1.81	0.11	0.51	0.618	1.83
L Thalamus [−16 −18 0]	−0.16	−0.96	0.342	2.52	−0.16	−0.69	0.499	1.92
R Thalamus [14 −14 2]	0.30	1.89	0.067	2.20	0.04	0.20	0.841	1.65
L preMotor [−16 −6 62]	−0.06	−0.34	0.739	2.75	−0.05	−0.22	0.824	1.80
L Insula [−38 −8 −10]	0.20	1.45	0.155	1.64	−0.12	−0.59	0.561	1.49
L MFG [−34 2 64]	0.01	0.05	0.960	1.76	−0.08	−0.30	0.768	2.38
R COper [42 −8 16]	0.17	1.21	0.232	1.70	−0.04	−0.16	0.871	1.82

Note: Regions of interest (ROIs) were defined from Schwartz et al. (2021). All β values indicate standardized weight for each ROI.

Abbreviations: COper, central operculum; HIPP, hippocampus; IPS, intraparietal sulcus; L, left; LOC, lateral occipital cortex; MFG, middle frontal gyrus; PreCG, precentral gyrus; preMotor, premotor cortex; R, right. * $p < 0.05$, *** $p < 0.001$.

$BF_{10} = 0.44$) and post-training ($r = 0.071$, $p = 0.743$, $BF_{10} = 0.46$), with no significant difference between correlations at each time point ($p = 0.830$). Together, these findings suggest the degree of association between non-symbolic and symbolic number comparison task performance remained similar following 4 weeks of no intervention.

3.5 | Training-related changes recapitulate neurodevelopmental shift in the association between cross-format (non-symbolic–symbolic) neural representational similarity (NRS) and arithmetic fluency

Our next major goal of the study was to examine whether changes in the relation between cross-format NRS and arithmetic skills follow-

ing number sense training recapitulate the developmental differences between childhood and adolescence (Schwartz et al., 2021). Children's arithmetic skills, assessed by standardized scores of the WJ-III Math Fluency, were not significantly different between pre- and post-training in the STT sample ($p = 0.272$). WJ-III Math Fluency scores from each time point in the STT sample were also not significantly different from those of children and adolescents in the LTD sample ($ps > 0.806$).

To investigate whether training-induced changes in underlying neural mechanisms support children's arithmetic skills in a similar or different way from long-term development, we examined the brain regions that showed a significant relation between cross-format NRS and arithmetic fluency in children in the LTD sample (Schwartz et al. (2021); a total of 15 ROIs; see Methods). In the STT sample, all ROIs except for the right hippocampus ($p = 0.034$, $BF_{10} = 2.55$) showed

Short-term Training (STT) Sample

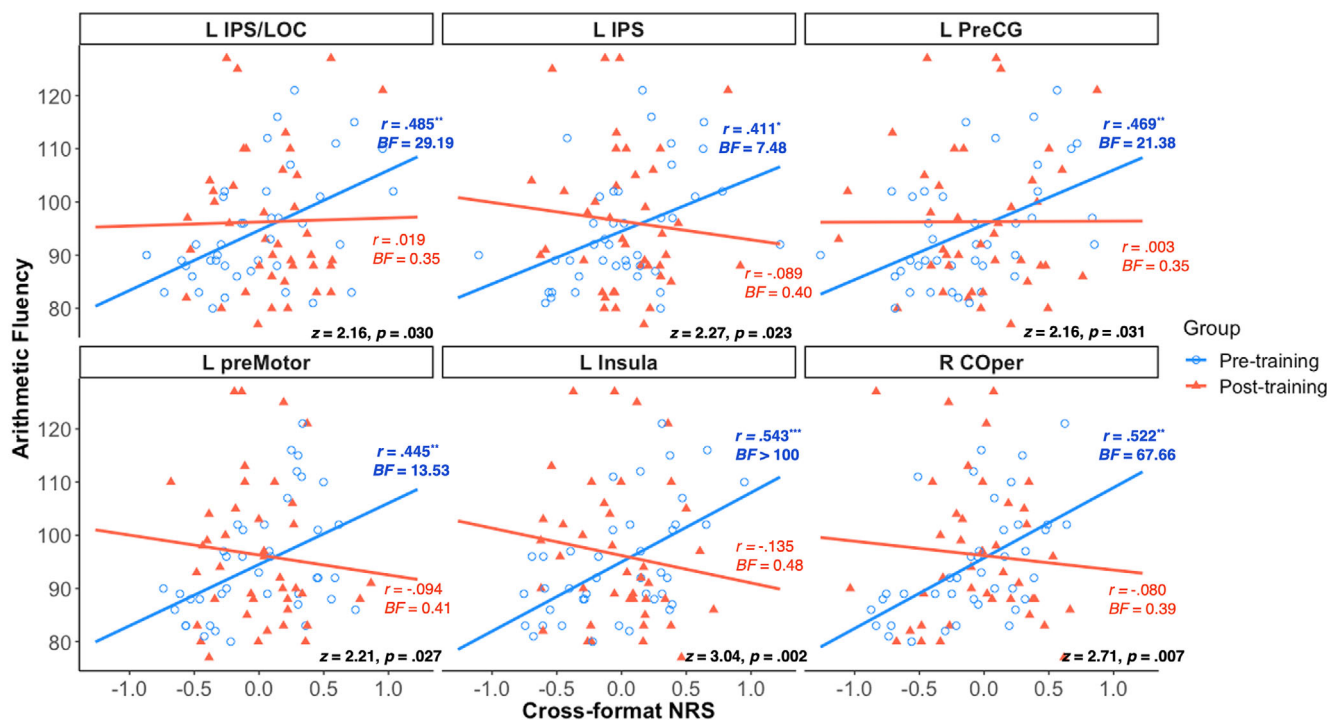


FIGURE 2 Short-term training leads to dissociation between cross-format (non-symbolic–symbolic) neural representational similarity (NRS) in parietal and frontal regions and arithmetic fluency in children. Before training (blue dots and regression lines), cross-format NRS significantly correlates with arithmetic fluency in children in the left intraparietal sulcus/lateral occipital cortex (IPS/LOC), precentral gyrus (PreCG), premotor cortex (premotor), insula and the right central operculum (COper) (Bayes factors [BF_{10}] > 3). In contrast, after training (red triangles and regression lines), cross-format NRS does not significantly correlate with arithmetic fluency in children in these regions of interest (ROIs) (BF_{10} < 3). Correlation coefficients were significantly different between pre- and post-training in these six ROIs (p s < 0.04), which were among the ROIs defined from Schwartz et al. (2021) (see also Figure 4a). Results from the other nine ROIs also showed significant brain-behavior correlations before, but not after, short-term training (Supplementary Results). p^* < 0.05, p^{**} < 0.01, p^{***} < 0.001. Statistically significant r -values and their corresponding p -values, as well as BF_{10} values that provide evidence for test hypothesis are bolded. L, left; R, right.

significant positive correlation between cross-format NRS and arithmetic fluency (r s > 0.34, p s < 0.026, BF_{10} s > 3.14) before training, replicating the results of children in the LTD sample (r s > 0.35, p s < 0.010, BF_{10} s > 6.71; Table 2). In contrast, after training, the same children in the STT sample did not show a significant relation between cross-format NRS and arithmetic fluency in any of the ROIs (r s < 0.33, p s > 0.043, BF_{10} s < 2.15; only L MGF showed p < 0.05 but with BF_{10} < 3), similar to the results observed with adolescents in the LTD sample (r s < 0.27, p s > 0.083, BF_{10} s < 1.28). Furthermore, a significant difference in correlation between cross-format NRS and arithmetic fluency was observed, across time points, in the STT sample in the left insula and right central operculum (difference in correlation coefficient: z s > 2.71, p s < 0.007, BF_{10} s > 4.09), although parietal cortical regions, including the left IPS, precentral gyrus, and premotor cortex, showed differences only in terms of p -values (difference in correlation coefficient: z s > 2.16, p s < 0.032) (Figure 2; see Table 2).

We then compared the relation between cross-format NRS and arithmetic fluency in the STT sample prior to training and children from the LTD sample. We found that the correlation between cross-format NRS and fluency in children at pre-training in the STT sample

was not significantly different from that of children in the LTD sample (difference in correlation coefficient: z s < 0.87, p s > 0.391, BF_{10} s < 0.28; Table 3). Furthermore, the correlation in children in the STT sample post-training was not significantly different from that of adolescents in the LTD sample (difference in correlation coefficient: z s < 1.50, p s > 0.135, BF_{10} s < 0.76; Table 4; see Figure 3 for results from ROIs plotted in Figure 2). These results demonstrate that the patterns of cross-format NRS-fluency relation in children at pre- and post-training mirror those observed in children and adolescents, respectively.

To confirm and extend these findings, we performed a multilinear regression analysis, which examined whether cross-format NRS in all ROIs jointly predict arithmetic fluency at each time point in the STT sample or each group in the LTD sample. We found that the model predicting arithmetic fluency from cross-format NRS in the ROIs was significant before training ($R^2_{Adjusted}$ = 0.486, p = 0.003) but not after training ($R^2_{Adjusted}$ = -0.039, p = 0.572) in the STT sample (Figure 4 and Table 5). In an analogous multilinear regression analysis with the LTD sample, the results mirrored those of the STT sample with the model being significant in children ($R^2_{Adjusted}$ = 0.423, p = 0.0009) but not in adolescents ($R^2_{Adjusted}$ = -0.191, p = 0.894) (Figure 4 and Table 6).

Short-term Training (STT) vs. Long-term Developmental (LTD) Sample

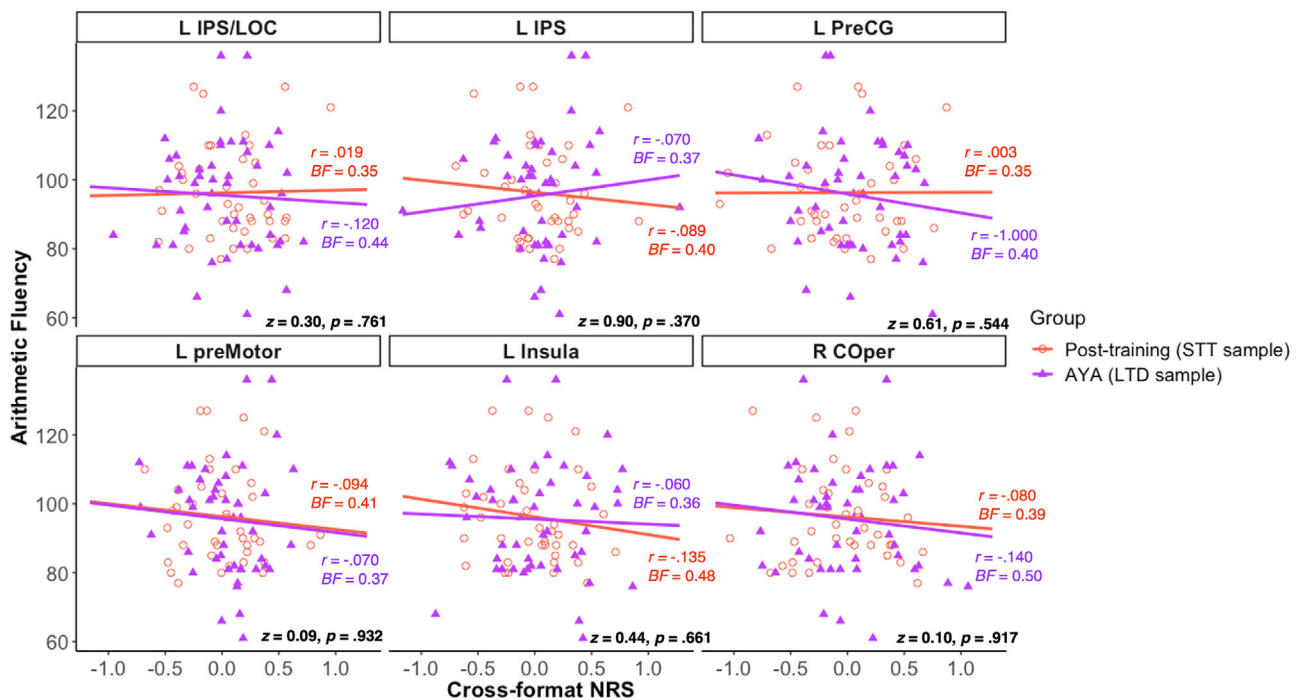


FIGURE 3 Association between cross-format (non-symbolic-symbolic) neural representational similarity (NRS) and arithmetic fluency in children after training is comparable to that of adolescents from a long-term developmental (LTD) sample. Cross-format NRS of children at post-training from the short-term training (STT) sample (red triangles and regression lines) and that of adolescents group from the LTD sample (purple triangles and regression lines; see details in Figure 1 and Methods) do not significantly correlate with arithmetic fluency in the left intraparietal sulcus/lateral occipital cortex (IPS/LOC), precentral gyrus (PreCG), premotor cortex (premotor), insula and the right central operculum (COper) (Bayes factors [BF_{10}] < 3). Correlation coefficients were not significantly different between children at post-training from STT and adolescents from LTD in these 6 ROIs (p s > 0.36), which were among the ROIs defined from Schwartz et al. (2021) (see also Figure 4a). Results from the other nine ROIs also showed no significant brain-behavior correlations at post-training in children in the STT sample and in adolescents in the LTD sample (Supplementary Results). L, left; R, right.

These results suggest that training-related changes in multivariate brain-behavior relation between cross-format NRS and arithmetic fluency were comparable to changes observed to occur across early-to-late developmental stage. Together, these results indicate that a short-term number sense training in early childhood may facilitate children's development towards dissociation between cross-format neural mapping and arithmetic fluency.

4 | DISCUSSION

We implemented a 4-week number sense intervention with the aim of enhancing neural mapping across non-symbolic and symbolic quantity representations in early elementary school children. Our primary objective was to test the hypothesis that STT-induced neural plasticity in children could recapitulate long-term neurodevelopmental changes observed from childhood to adolescence. The results of our study demonstrated several significant findings. First, we found that short-term number sense training improved performance on both non-symbolic and symbolic number comparison tasks. Second, number

sense training induced a dissociation between non-symbolic and symbolic quantity representations at the behavioral level. Third, number sense training altered the relation between cross-format neural representational similarity (NRS) and arithmetic fluency. Remarkably, these training-induced changes mirrored developmental changes observed from childhood to adolescence. Taken together, these findings provide compelling evidence that intensive STT in children's number sense can recapitulate neurocognitive changes across development.

Our training program encompassed both conceptual and procedural aspects of number knowledge, targeting key principles of numerical quantity such as establishing one-to-one correspondence between non-symbolic and symbolic numbers and understanding cardinality and ordinality of quantity. Additionally, the program included extensive practice of number comparison tasks across different formats, with increasing emphasis on symbolic numbers over the course of training. The results of our study indicated that our number sense training had a significant impact on improving the performance efficiency of both non-symbolic and symbolic number comparison tasks. Further analysis revealed that observed improvements in efficiency were primarily driven by changes in RTs for

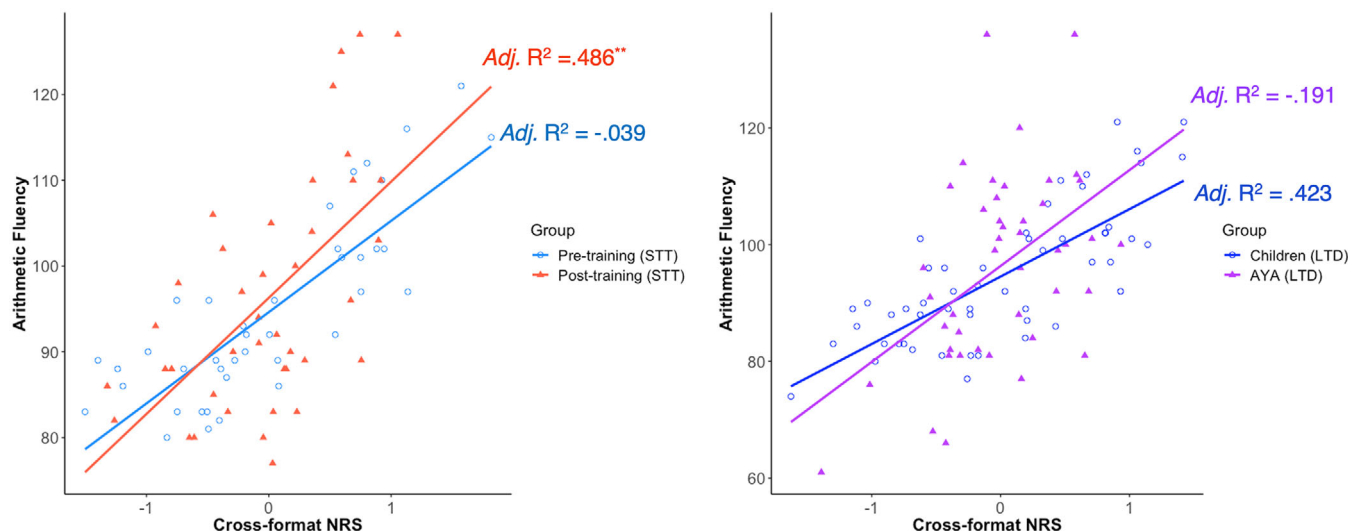


FIGURE 4 Cross-format (non-symbolic–symbolic) neural representational similarity (NRS) in multiple brain regions jointly contributes to arithmetic fluency in children before, but not after, short-term training. (a) Regions of interest (ROIs). A total of 15 ROIs were defined from Schwartz et al. (2021). (b) Short-term training (STT) sample. Cross-format NRS in all ROIs jointly significantly predicted arithmetic fluency in children at pre-training (adjusted $R^2 = 0.486$, $p = 0.003$), but not at post-training (adjusted $R^2 = -0.039$, $p = 0.572$). (c) Long-term developmental (LTD) sample. Cross-format NRS in all ROIs jointly significantly predicted arithmetic fluency in children (adjusted $R^2 = 0.423$, $p = 0.003$), but not in adolescents (adjusted $R^2 = -0.191$, $p = 0.894$). Each adjusted R^2 value was obtained from a multilinear regression analysis predicting arithmetic fluency with cross-format NRS values of all 15 ROIs in each group of each sample included as predictors (see details in Methods). $p^{**} < 0.01$, $p^{***} < 0.001$. COper, central operculum; HIPp, hippocampus; IPS, intraparietal sulcus; L, left; MFG, middle frontal gyrus; PreCG, precentral gyrus; preMotor, premotor cortex; R, right.

both non-symbolic and symbolic comparison tasks. Importantly, these changes were specifically observed in the training group and not in the control group, underscoring the effectiveness of our intervention. Taken together, these results provide evidence that our number sense training program enhanced children's performance on number comparison tasks, aligning with prior research on the effectiveness of similar interventions (e.g., Honore & Noel, 2016; Kucian et al., 2011; Obersteiner et al., 2013). More generally, our findings suggest that a comprehensive approach to children's learning, including a combination of conceptual and procedural learning, may be effective in number sense development.

Next, we investigated the impact of number sense training on the relationship between NDEs observed in non-symbolic and symbolic number comparison tasks. NDE, which assesses the discriminability between numerical quantities (Buckley & Gillman, 1974; Moyer & Landauer, 1967), has been thought to provide a measure of the sensitivity to differences in quantity independent of overall performance (Brannon & Merritt, 2011; Kaufmann et al., 2008; Nieder & Dehaene, 2009). In a previous study, NDE on non-symbolic and symbolic number comparison tasks were significantly related to each other in children but not in adolescents (Schwartz et al., 2021). The observed pattern of a shift from overlap to dissociation between number formats is consistent with the developmental estrangement account (Lyons et al., 2012; Nakai et al., 2023; Schwartz et al., 2021). Similarly, in the current training study, we observed that prior to training, there was a significant relation in NDE between non-symbolic and symbolic quantities; following training, this significant association no longer persisted,

indicating dissociation in quantity representations at the behavioral level. Notably, we did not observe such a training-induced dissociation in overall efficiency scores, as their association between formats remained significant both before and after training.

Our findings of the training-induced shift from an association to dissociation between number formats, which were specifically observed for NDE and not overall efficiency scores, suggest that NDE may be sensitive to precise quantity representations. In contrast, efficiency scores may more likely reflect processing efficiency associated with other cognitive processes such as processing speed or decision-making. Our findings indicate that STT can facilitate the development of more distinct representations of quantity between non-symbolic and symbolic formats, providing support for the estrangement account (Bulthé et al., 2018; Bulthé et al., 2019; Lyons et al., 2012; Wilkey et al., 2020). The observed cross-format dissociation in NDE after training also provides insights into how our training may have improved distinct aspects of numerical processing. Our number sense training protocol was structured in a way that gradually progressed from mapping between non-symbolic and symbolic quantities to fluency of symbolic numbers. Such semantic associations between non-symbolic and symbolic numbers may solidify conceptual understanding of symbolic numbers and facilitate a dissociation between non-symbolic and symbolic number formats upon mastery of symbolic number knowledge. Our findings underscore the potential of short-term interventions in shaping cognitive representations of numerical quantity and highlight the parallels between short-term learning and long-term learning and development.



Critically, our findings from multivariate neural pattern analysis recapitulated long-term neurodevelopmental changes in relations between cross-format NRS and arithmetic skills. Across the STT and LTD samples, we found that cross-format NRS in distributed parietal and frontal cortical regions is significantly correlated with arithmetic fluency before training or at early developmental stage, and that this association becomes non-significant following training or at the late developmental stage. These patterns of findings observed in both the STT and LTD samples were evident across individual regions in whole brain analyses and distributed regions in multivariate prediction models. Furthermore, associations between cross-format NRS and arithmetic skills at post-training in the STT sample were not significantly different from those of adolescents in the LTD sample. Together, our study demonstrates that training-induced learning can alter the association between cross-format NRS and arithmetic fluency in key brain regions associated with numerical cognition, similar to learning that occurs across long-term development.

Our findings provide further support for, and extend, the estrangement account, which is broadly consistent with the notion that later in development, arithmetic fluency would no longer depend on the association between number formats. Previous studies have even shown that the degree of neural associations between number formats is *negatively* correlated with arithmetic skills in adults (Bulthé et al., 2018; Bulthé et al., 2019; Wilkey et al., 2020). These findings are also consistent with the notion that functional circuits or representations supporting numerical magnitude processing in different formats become “decoupled” over development (Nakai et al., 2023; Skagenholt et al., 2022) and engage different coding mechanisms in adulthood (Bulthé et al., 2014; Lyons, Ansari, & Beilock, 2015; Lyons & Beilock, 2018; Sokolowski et al., 2019; Yeo et al., 2020).

Overall, these results reveal that targeted number sense intervention induces significant changes in the association between cross-format NRS and arithmetic problem-solving abilities in children, which, after training were indistinguishable from the pattern of associations observed in adolescents (Schwartz et al., 2021). Crucially, our findings support the notion that short-term learning experiences can recapitulate long-term neurodevelopmental changes observed in children’s cognitive development, aligning with previous research demonstrating a similar recapitulation of longitudinal brain response changes in the context of arithmetic training (Rosenberg-Lee et al., 2018).

More broadly, our findings are consistent with the idea that interventions can enhance experience-dependent neural plasticity crucial for adaptive learning (Fandakova & Hartley, 2020; Galván, 2010; May, 2011; McLaughlin et al., 2018; Nelson, 2000). Targeted intervention studies have the potential to address the impact of targeted educational experiences on neural plasticity (Cooper & Mackey, 2016; Cramer et al., 2011; Johnson, 2001, 2011; Liu et al., 2023; Park & Mackey, 2022; Perdue et al., 2022; Wenger & Lövdén, 2016; Weyandt et al., 2020). Such interventions can induce changes in neural representations or circuits associated with a specific skillset, leading to more efficient information processing (Johnson, 2001, 2011). By demonstrating how a targeted training protocol can induce changes in neural representational patterns associated with behavior in children, our

study contributes to a mechanistic understanding of how children acquire numerical abilities and highlights the importance of examining neural representational patterns associated with cognitive skill development in early childhood. Moreover, our approach presents valuable opportunities to explore the effects of neuromodulation techniques, such as transcranial direct current stimulation (TDCS) (Kadosh, 2014; Kadosh et al., 2010; Looi et al., 2017; Sarkar & Cohen Kadosh, 2016) on neural representations pertinent to enhancing cognitive abilities in neurodiverse populations.

4.1 | Limitations and future work

Across all brain ROIs derived from the LTD sample, our STT recapitulated long-term neurodevelopmental changes in the relation between cross-format NRS and arithmetic fluency. However, not all regions exhibited significant differences in this relationship from pre- to post-training. To better understand the specific factors contributing to these variations, future studies should examine behavioral and brain measures at multiple time points from larger cohorts of children (Schönbrodt & Perugini, 2013). Moreover, studies integrating active control interventions are necessary to gain a deeper understanding and pinpoint the mechanisms accountable for the observed training effects in the current study.

We employed NDE as a behavioral metric of quantity discriminability, aiming for consistency with the LTD sample (Schwartz et al., 2021). While NDE has been widely used for this purpose, some research suggests that its applicability to symbolic numbers may not exclusively reflect numerical quantity representation or processing. For example, previous studies suggest that NDE may also reflect general cognitive processes such as attention and response selection (Lyons, Nuerk, & Ansari, 2015; Van Opstal et al., 2009; Van Opstal et al., 2008). Further studies are needed to validate the robustness of our findings and to explore alternative measures of quantity processing. Lastly, in light of differences between the processing of small and large numbers (Hutchison et al., 2020; Lyons & Beilock, 2018), it will be important to investigate how NRS evolves with varying numerical magnitudes of quantity. This inquiry could shed light on the generalizability of neural processing mechanisms across different quantity ranges and numerical formats.

5 | CONCLUSION

Our study reveals that short-term, targeted training in number sense can replicate developmental changes traditionally seen over prolonged periods. We observed a dissociation between neural mappings of quantity formats and problem-solving abilities in key cortical regions after just 4 weeks of training, which mirrored changes that LTD changes that occur across early-to-late developmental stages. This underscores the transformative potential of early, focused educational interventions. Utilizing multivariate pattern analysis, our research offers a more mechanistic view into how learning reshapes neural circuits,

advancing our understanding of how cognitive skills and acquired and refined. Our methodological innovations enabled a deeper exploration of the mechanisms underlying learning and cognitive development, and highlighted the potential of neural representation analysis in enhancing our understanding of learning and development. Our findings underscore the significant impact of early, tailored training on cognitive development, opening new avenues for educational practices aimed at individual learning needs.

ACKNOWLEDGMENTS

This research was supported by grants from the National Institutes of Health (HD094623, HD059205, MH084164) and National Science Foundation (DRL-2024856) to V.M. and Stanford Maternal & Child Health Research Institute Postdoctoral Support Award to H.C. We thank Kristen Pilner Blair for assistance with game design and participating families for assistance with the study.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interest.

DATA AVAILABILITY STATEMENT

Data that support the findings of this study are available upon request from corresponding authors.

ORCID

Yunji Park  <https://orcid.org/0000-0003-2904-9505>

Yuan Zhang  <https://orcid.org/0000-0001-9157-5544>

Hyesang Chang  <https://orcid.org/0000-0002-2231-1112>

REFERENCES

- Ansari, D. (2008). Effects of development and enculturation on number representation in the brain. *Nature Reviews Neuroscience*, 9, 278–291. <https://doi.org/10.1038/nrn2334>
- Ansari, D., & Dhital, B. (2006). Age-related changes in the activation of the intraparietal sulcus during nonsymbolic magnitude processing: An event-related functional magnetic resonance imaging study. *Journal of Cognitive Neuroscience*, 18(11), 1820–1828. <https://doi.org/10.1162/jocn.2006.18.11.1820>
- Arsalidou, M., Pawliw-Levac, M., Sadeghi, M., & Pascual-Leone, J. (2018). Brain areas associated with numbers and calculations in children: Meta-analyses of fMRI studies. *Developmental Cognitive Neuroscience*, 30, 239–250. <https://doi.org/10.1016/j.dcn.2017.08.002>
- Ashburner, J., Barnes, G., Chen, C.-C., Daunizeau, J., Flandin, G., Friston, K., Kiebel, S., Kilner, J., Litvak, V., & Moran, R. (2014). *SPM12 manual* (Vol. 2464, pp. 4). Wellcome Trust Centre for Neuroimaging.
- Ashkenazi, S., Rosenberg-Lee, M., Metcalfe, A. W., Swigart, A. G., & Menon, V. (2013). Visuo-spatial working memory is an important source of domain-general vulnerability in the development of arithmetic cognition. *Neuropsychologia*, 51(11), 2305–2317. <https://doi.org/10.1016/j.neuropsychologia.2013.06.031>
- Bakeman, R. (2005). Recommended effect size statistics for repeated measures designs. *Behavior Research Methods*, 37(3), 379–384. <https://doi.org/10.3758/BF03192707>
- Blair, K. P. (2013a, April). Feedback in critter corral: The effectiveness of implication versus corrective feedback in a math learning game. In 2013 Early education and technology for children conference, Salt Lake City, UT.
- Blair, K. P. (2013b). Learning in critter corral: Evaluating three kinds of feedback in a preschool math game. In Proceedings of the interaction design and children 2013 conference, ACM, New York.
- Brannon, E. M., & Merritt, D. J. (2011). Evolutionary foundations of the approximate number system. *Space, Time and Number in the Brain*, 207–224. <https://doi.org/10.1016/B978-0-12-385948-8.00014-1>
- Bruyer, R., & Brysbaert, M. (2011). Combining speed and accuracy in cognitive psychology: Is the inverse efficiency score (IES) a better dependent variable than the mean reaction time (RT) and the percentage of errors (PE)? *Psychologica Belgica*, 51(1), 5–13. <https://doi.org/10.5334/pb-51-1-5>
- Bulthé, J., De Smedt, B., & de Beeck, H. O. (2014). Format-dependent representations of symbolic and non-symbolic numbers in the human cortex as revealed by multi-voxel pattern analyses. *Neuroimage*, 87, 311–322. <https://doi.org/10.1016/j.neuroimage.2013.10.049>
- Bulthé, J., De Smedt, B., & Op de Beeck, H. (2018). Arithmetic skills correlate negatively with the overlap of symbolic and non-symbolic number representations in the brain. *Cortex*, 101, 306–308. <https://doi.org/10.1016/j.cortex.2018.01.008>
- Bulthé, J., Prinsen, J., Vanderauwera, J., Duyck, S., Daniels, N., Gillebert, C. R., Mantini, D., Op de Beeck, H. P., & De Smedt, B. (2019). Multi-method brain imaging reveals impaired representations of number as well as altered connectivity in adults with dyscalculia. *Neuroimage*, 190, 289–302. <https://doi.org/10.1016/j.neuroimage.2018.06.012>
- Burgund, E. D., Kang, H. C., Kelly, J. E., Buckner, R. L., Snyder, A. Z., Petersen, S. E., & Schlaggar, B. L. (2002). The feasibility of a common stereotactic space for children and adults in fMRI studies of development. *Neuroimage*, 17(1), 184–200. <https://doi.org/10.1006/nimg.2002.1174>
- Butterworth, B., & Walsh, V. (2011). Neural basis of mathematical cognition. *Current Biology*, 21(16), R618–R621. <https://doi.org/10.1016/j.cub.2011.07.005>
- Cantlon, J. F., Brannon, E. M., Carter, E. J., & Pelphrey, K. A. (2006). Functional imaging of numerical processing in adults and 4-y-old children. *PLoS Biology*, 4(5), e125. <https://doi.org/10.1371/journal.pbio.0040125>
- Chang, H., Chen, L., Zhang, Y., Xie, Y., de los Angeles, C., Adair, E., Zanitti, G., Wassermann, D., Rosenberg-Lee, M., & Menon, V. (2022). Foundational number sense training gains are predicted by hippocampal-parietal circuits. *Journal of Neuroscience*, 42(19), 4000–4015.
- Chang, H., Rosenberg-Lee, M., Qin, S., & Menon, V. (2019). Faster learners transfer their knowledge better: Behavioral, mnemonic, and neural mechanisms of individual differences in children's learning. *Developmental Cognitive Neuroscience*, 40, 100719. <https://doi.org/10.1016/j.dcn.2019.100719>
- Cohen, J. (1992). Statistical power analysis. *Current Directions in Psychological Science*, 1(3), 98–101. <https://doi.org/10.1111/1467-8721.ep10768783>
- Cooper, E. A., & Mackey, A. P. (2016). Sensory and cognitive plasticity: Implications for academic interventions. *Current Opinion in Behavioral Sciences*, 10, 21–27. <https://doi.org/10.1016/j.cobeha.2016.04.008>
- Cramer, S. C., Sur, M., Dobkin, B. H., O'Brien, C., Sanger, T. D., Trojanowski, J. Q., Rumsey, J. M., Hicks, R., Cameron, J., & Chen, D. (2011). Harnessing neuroplasticity for clinical applications. *Brain*, 134(6), 1591–1609. <https://doi.org/10.1093/brain/awr039>
- Fandakova, Y., & Hartley, C. A. (2020). Mechanisms of learning and plasticity in childhood and adolescence. *Developmental Cognitive Neuroscience*, 42, 100764. <https://doi.org/10.1016/j.dcn.2020.100764>
- Faulkenberry, T. J., Ly, A., & Wagenmakers, E.-J. (2020). Bayesian inference in numerical cognition: A tutorial using JASP. *Journal of Numerical Cognition*, 6(2), 231–259. <https://doi.org/10.5964/jnc.v6i2.288>
- Galván, A. (2010). Neural plasticity of development and learning. *Human Brain Mapping*, 31(6), 879–890. <https://doi.org/10.1002/hbm.21029>
- Gillman, C. B. (1974). Comparisons of digits and dot patterns. *Journal of Experimental Psychology*, 103(6), 1131–1136.
- Gilmore, C., Attridge, N., Clayton, S., Cragg, L., Johnson, S., Marlow, N., Simms, V., & Inglis, M. (2013). Individual differences in inhibitory control,



- not non-verbal number acuity, correlate with mathematics achievement. *PLoS One*, 8(6), e67374. <https://doi.org/10.1371/journal.pone.0067374>
- Glover, G. H., & Lai, S. (1998). Self-navigated spiral fMRI: interleaved versus single-shot. *Magnetic Resonance in Medicine*, 39(3), 361–368.
- Holloway, I. D., & Ansari, D. (2009). Mapping numerical magnitudes onto symbols: The numerical distance effect and individual differences in children's mathematics achievement. *Journal of Experimental Child Psychology*, 103(1), 17–29. <https://doi.org/10.1016/j.jecp.2008.04.001>
- Honore, N., & Noel, M. P. (2016). Improving preschoolers' arithmetic through number magnitude training: The impact of non-symbolic and symbolic training. *PLoS One*, 11(11), e0166685. <https://doi.org/10.1371/journal.pone.0166685>
- Hutchison, J. E., Ansari, D., Zheng, S., De Jesus, S., & Lyons, I. M. (2020). The relation between subitizable symbolic and non-symbolic number processing over the kindergarten school year. *Developmental Science*, 23(2), e12884. <https://doi.org/10.1111/desc.12884>
- Iuculano, T., Rosenberg-Lee, M., Richardson, J., Tenison, C., Fuchs, L., Supekar, K., & Menon, V. (2015). Cognitive tutoring induces widespread neuroplasticity and remediates brain function in children with mathematical learning disabilities. *Nature Communications*, 6, 8453. <https://doi.org/10.1038/ncomms9453>
- Jeffreys, H. (1961). *The theory of probability*. Oxford University Press.
- Johnson, M. H. (2001). Functional brain development in humans. *Nature Reviews Neuroscience*, 2(7), 475–483. <https://doi.org/10.1038/35081509>
- Johnson, M. H. (2011). Interactive specialization: A domain-general framework for human functional brain development? *Developmental Cognitive Neuroscience*, 1(1), 7–21. <https://doi.org/10.1016/j.dcn.2010.07.003>
- Jordan, N. C., Devlin, B. L., & Botello, M. (2022). Core foundations of early mathematics: Refining the number sense framework. *Current Opinion in Behavioral Sciences*, 46, 101181. <https://doi.org/10.1016/j.cobeha.2022.101181>
- Jordan, N. C., Kaplan, D., Ramineni, C., & Locuniak, M. N. (2009). Early math matters: Kindergarten number competence and later mathematics outcomes. *Developmental Psychology*, 45(3), 850. <https://doi.org/10.1037/a0014939>
- Kadosh, R. C. (2014). *The stimulated brain: Cognitive enhancement using non-invasive brain stimulation*. Elsevier.
- Kadosh, R. C., Soskic, S., Iuculano, T., Kanai, R., & Walsh, V. (2010). Modulating neuronal activity produces specific and long-lasting changes in numerical competence. *Current Biology*, 20(22), 2016–2020. <https://doi.org/10.1016/j.cub.2010.10.007>
- Kang, H. C., Burgund, E. D., Lugar, H. M., Petersen, S. E., & Schlaggar, B. L. (2003). Comparison of functional activation foci in children and adults using a common stereotactic space. *Neuroimage*, 19(1), 16–28. [https://doi.org/10.1016/S1053-8119\(03\)00038-7](https://doi.org/10.1016/S1053-8119(03)00038-7)
- Karraker, S., Schwartz, F., Chang, H., Iuculano, T., Adair, E., Rosenberg-Lee, M., Chen, L., & Menon, V. (2018). Number training remediates math difficulties in children: behavioral and neural correlates. In 6th Biennial meeting of the International Mind, Brain, and Education Society, Los Angeles, CA.
- Kaufmann, L., Vogel, S. E., Wood, G., Kremser, C., Schocke, M., Zimmerhackl, L.-B., & Koenig, J. W. (2008). A developmental fMRI study of nonsymbolic numerical and spatial processing. *Cortex*, 44(4), 376–385. <https://doi.org/10.1016/j.cortex.2007.08.003>
- Kersey, A. J., & Cantlon, J. F. (2017). Neural tuning to numerosity relates to perceptual tuning in 3–6-year-old children. *Journal of Neuroscience*, 37(3), 512–522. <https://doi.org/10.1523/JNEUROSCI.0065-16.2016>
- Keyesers, C., Gazzola, V., & Wagenmakers, E.-J. (2020). Using Bayes factor hypothesis testing in neuroscience to establish evidence of absence. *Nature Neuroscience*, 23(7), 788–799. <https://doi.org/10.1038/s41593-020-0660-4>
- Kriegeskorte, N., Goebel, R., & Bandettini, P. (2006). Information-based functional brain mapping. *Proceedings of the National Academy of Sciences of the United States of America*, 103(10), 3863. <https://doi.org/10.1073/pnas.0600244103>
- Kucian, K., Grond, U., Rotzer, S., Henzi, B., Schonmann, C., Plangger, F., Galli, M., Martin, E., & von Aster, M. (2011). Mental number line training in children with developmental dyscalculia. *Neuroimage*, 57(3), 782–795. <https://doi.org/10.1016/j.neuroimage.2011.01.070>
- Lau, N. T., Merkley, R., Tremblay, P., Zhang, S., De Jesus, S., & Ansari, D. (2021). Kindergarteners' symbolic number abilities predict nonsymbolic number abilities and math achievement in grade 1. *Developmental Psychology*, 57(4), 471. <https://doi.org/10.1037/dev0001158>
- Liu, J., Chang, H., Abrams, D. A., Kang, J. B., Chen, L., Rosenberg-Lee, M., & Menon, V. (2023). Atypical cognitive training-induced learning and brain plasticity and their relation to insistence on sameness in children with autism. *Elife*, 12, e86035. <https://doi.org/10.7554/eLife.86035>
- Looi, C. Y., Lim, J., Sella, F., Lollot, S., Duta, M., Avramenko, A. A., & Cohen Kadosh, R. (2017). Transcranial random noise stimulation and cognitive training to improve learning and cognition of the atypically developing brain: A pilot study. *Scientific Reports*, 7(1), 4633. <https://doi.org/10.1038/s41598-017-04649-x>
- Lyons, I. M., Ansari, D., & Beilock, S. L. (2012). Symbolic estrangement: Evidence against a strong association between numerical symbols and the quantities they represent. *Journal of Experimental Psychology General*, 141(4), 635–641. <https://doi.org/10.1037/a0027248>
- Lyons, I. M., Ansari, D., & Beilock, S. L. (2015). Qualitatively different coding of symbolic and nonsymbolic numbers in the human brain. *Human Brain Mapping*, 36(2), 475–488. <https://doi.org/10.1002/hbm.22641>
- Lyons, I. M., & Beilock, S. L. (2018). Characterizing the neural coding of symbolic quantities. *Neuroimage*, 178, 503–518. <https://doi.org/10.1016/j.neuroimage.2018.05.062>
- Lyons, I. M., Nuerk, H.-C., & Ansari, D. (2015). Rethinking the implications of numerical ratio effects for understanding the development of representational precision and numerical processing across formats. *Journal of Experimental Psychology: General*, 144(5), 1021. <https://doi.org/10.1037/xge0000094>
- May, A. (2011). Experience-dependent structural plasticity in the adult human brain. *Trends in Cognitive Sciences*, 15(10), 475–482. <https://doi.org/10.1016/j.tics.2011.08.002>
- McLaughlin, K. A., Mackey, A., Fernandes, G. F., Brown, K., Bühler, J. C., & Bunge, S. A. (2018). Human brain plasticity: Future research directions and implications for children's learning and development. *Jacobs Foundation* (pp. 1–18).
- Michels, L., O'Gorman, R., & Kucian, K. (2018). Functional hyperconnectivity vanishes in children with developmental dyscalculia after numerical intervention. *Developmental Cognitive Neuroscience*, 30, 291–303. <https://doi.org/10.1016/j.dcn.2017.03.005>
- Moyer, R. S., & Landauer, T. K. (1967). Time required for judgements of numerical inequality. *Nature*, 215(5109), 1519–1520. <https://doi.org/10.1038/2151519a0>
- Nakai, T., Girard, C., Longo, L., Chesnokova, H., & Prado, J. (2023). Cortical representations of numbers and nonsymbolic quantities expand and segregate in children from 5 to 8 years of age. *PLoS Biology*, 21(1), e3001935. <https://doi.org/10.1371/journal.pbio.3001935>
- Nelson, C. A. (2000). Neural plasticity and human development: The role of early experience in sculpting memory systems. *Developmental Science*, 3(2), 115–136. <https://doi.org/10.1111/1467-7687.00104>
- Nieder, A., & Dehaene, S. (2009). Representation of number in the brain. *Annual Review of Neuroscience*, 32, 185–208. <https://doi.org/10.1146/annurev.neuro.051508.135550>
- Obersteiner, A., Reiss, K., & Ufer, S. (2013). How training on exact or approximate mental representations of number can enhance first-grade students' basic number processing and arithmetic skills. *Learning and Instruction*, 23, 125–135. <https://doi.org/10.1016/j.learninstruc.2012.08.004>
- Park, A. T., & Mackey, A. P. (2022). Do younger children benefit more from cognitive and academic interventions? How training studies can

- provide insights into developmental changes in plasticity. *Mind, Brain, and Education*, 16(1), 24–35. <https://doi.org/10.1111/mbe.12304>
- Perdue, M. V., Mahaffy, K., Vlahcevic, K., Wolfman, E., Erbeli, F., Richlan, F., & Landi, N. (2022). Reading intervention and neuroplasticity: a systematic review and meta-analysis of brain changes associated with reading intervention. *Neuroscience & Biobehavioral Reviews*, 132, 465–494.
- Perrachione, T. K., Del Tufo, S. N., Winter, R., Murtagh, J., Cyr, A., Chang, P., Halverson, K., Ghosh, S. S., Christodoulou, J. A., & Gabrieli, J. D. (2016). Dysfunction of rapid neural adaptation in dyslexia. *Neuron*, 92(6), 1383–1397. <https://doi.org/10.1016/j.neuron.2016.11.020>
- Rosenberg-Lee, M., Iuculano, T., Bae, S. R., Richardson, J., Qin, S., Jolles, D., & Menon, V. (2018). Short-term cognitive training recapitulates hippocampal functional changes associated with one year of longitudinal skill development. *Trends in Neuroscience and Education*, 10, 19–29. <https://doi.org/10.1016/j.tine.2017.12.001>
- Rousselle, L., & Noel, M. P. (2007). Basic numerical skills in children with mathematics learning disabilities: a comparison of symbolic vs non-symbolic number magnitude processing. *Cognition*, 102(3), 361–395. <https://doi.org/10.1016/j.cognition.2006.01.005>
- Sarkar, A., & Cohen Kadosh, R. (2016). Transcranial electrical stimulation and numerical cognition. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale*, 70(1), 41. <https://doi.org/10.1037/cep0000064>
- Schneider, W., Eschman, A., & Zuccolotto, A. (2002). *E-Prime: User's guide*. Psychology Software Inc.
- Schönbrodt, F. D., & Perugini, M. (2013). At what sample size do correlations stabilize? *Journal of Research in Personality*, 47(5), 609–612. <https://doi.org/10.1016/j.jrp.2013.05.009>
- Schwartz, F., Zhang, Y., Chang, H., Karraker, S., Kang, J. B., & Menon, V. (2021). Neural representational similarity between symbolic and non-symbolic quantities predicts arithmetic skills in childhood but not adolescence. *Developmental Science*, 24(6), e13123. <https://doi.org/10.1111/desc.13123>
- Siegler, R. S., Duncan, G. J., Davis-Kean, P. E., Duckworth, K., Claessens, A., Engel, M., Susperreguy, M. I., & Chen, M. (2012). Early predictors of high school mathematics achievement. *Psychological Science*, 23(7), 691–697. <https://doi.org/10.1177/0956797612440101>
- Skagenholt, M., Skagerlund, K., & Träff, U. (2022). Neurodevelopmental differences in task-evoked number network connectivity: Comparing symbolic and nonsymbolic number discrimination in children and adults. *Developmental Cognitive Neuroscience*, 58, 101159. <https://doi.org/10.1016/j.dcn.2022.101159>
- Sokolowski, H. M., Fias, W., Mousa, A., & Ansari, D. (2017). Common and distinct brain regions in both parietal and frontal cortex support symbolic and nonsymbolic number processing in humans: A functional neuroimaging meta-analysis. *Neuroimage*, 146, 376–394. <https://doi.org/10.1016/j.neuroimage.2016.10.028>
- Sokolowski, H. M., Hawes, Z., Peters, L., & Ansari, D. (2019). Symbols are special: An fMRI adaptation study of symbolic, nonsymbolic and non-numerical magnitude processing in the human brain. *Cerebral Cortex Communications*, 2(3). <https://doi.org/10.31234/osf.io/xw2fq>
- Suárez-Pellicioni, M., & Booth, J. R. (2018). Fluency in symbolic arithmetic refines the approximate number system in parietal cortex. *Human Brain Mapping*, 39(10), 3956–3971. <https://doi.org/10.1002/hbm.24223>
- Szaflarski, J. P., Altaye, M., Rajagopal, A., Eaton, K., Meng, X., Plante, E., & Holland, S. K. (2012). A 10-year longitudinal fMRI study of narrative comprehension in children and adolescents. *Neuroimage*, 63(3), 1188–1195. <https://doi.org/10.1016/j.neuroimage.2012.08.049>
- Tobia, V., Bonifacci, P., & Marzocchi, G. M. (2021). Symbolic versus non-symbolic training for improving early numeracy in preschoolers at risk of developing difficulties in mathematics. *Research in Developmental Disabilities*, 111, 103893. <https://doi.org/10.1016/j.ridd.2021.103893>
- Van Opstal, F., Fias, W., Peigneux, P., & Verguts, T. (2009). The neural representation of extensively trained ordered sequences. *Neuroimage*, 47(1), 367–375. <https://doi.org/10.1016/j.neuroimage.2009.04.035>
- Van Opstal, F., Gevers, W., De Moor, W., & Verguts, T. (2008). Dissecting the symbolic distance effect: Comparison and priming effects in numerical and nonnumerical orders. *Psychonomic Bulletin & Review*, 15, 419–425.
- Wenger, E., & Lövdén, M. (2016). The Learning Hippocampus: Education and experience-dependent plasticity. *Mind, Brain, and Education*, 10(3), 171–183. <https://doi.org/10.1111/mbe.12112>
- Weyandt, L. L., Clarkin, C. M., Holding, E. Z., May, S. E., Marraccini, M. E., Gudmundsdottir, B. G., Shepard, E., & Thompson, L. (2020). Neuroplasticity in children and adolescents in response to treatment intervention: A systematic review of the literature. *Clinical and Translational Neuroscience*, 4(2), 21. <https://doi.org/10.1177/2514183X20974231>
- Wilkey, E. D., Conrad, B. N., Yeo, D. J., & Price, G. R. (2020). Shared numerosity representations across formats and tasks revealed with 7 Tesla fMRI: Decoding, generalization, and individual differences in behavior. *Cerebral Cortex Communications*, 1(1), tgaa038. <https://doi.org/10.1093/texcom/tgaa038>
- Wilson, A. J., Revkin, S. K., Cohen, D., Cohen, L., & Dehaene, S. (2006). An open trial assessment of “The Number Race”, an adaptive computer game for remediation of dyscalculia. *Behavioral and Brain Functions*, 2, 20. <https://doi.org/10.1186/1744-9081-2-20>
- Yeo, D. J., Pollack, C., Merkley, R., Ansari, D., & Price, G. R. (2020). The “Inferior Temporal Numeral Area” distinguishes numerals from other character categories during passive viewing: A representational similarity analysis. *Neuroimage*, 214, 116716. <https://doi.org/10.1016/j.neuroimage.2020.116716>

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Park, Y., Zhang, Y., Chang, H., & Menon, V. (2024). Short-term number sense training recapitulates long-term neurodevelopmental changes from childhood to adolescence. *Developmental Science*, 27, e13524. <https://doi.org/10.1111/desc.13524>