Cognitive tutoring induces widespread neuroplasticity and remediates brain function in children with mathematical learning disabilities

Teresa Iuculano1, Miriam Rosenberg-Lee1,*, Jennifer Richardson1,*, Caitlin Tenison1, Lynn Fuchs2, Kaustubh Supekar1 & Vinod Menon1,3,4

Competency with numbers is essential in today’s society; yet, up to 20% of children exhibit moderate to severe mathematical learning disabilities (MLD). Behavioural intervention can be effective, but the neurobiological mechanisms underlying successful intervention are unknown. Here we demonstrate that eight weeks of 1:1 cognitive tutoring not only remediates poor performance in children with MLD, but also induces widespread changes in brain activity. Neuroplasticity manifests as normalization of aberrant functional responses in a distributed network of parietal, prefrontal and ventral temporal–occipital areas that support successful numerical problem solving, and is correlated with performance gains. Remarkably, machine learning algorithms show that brain activity patterns in children with MLD are significantly discriminable from neurotypical peers before, but not after, tutoring, suggesting that behavioural gains are not due to compensatory mechanisms. Our study identifies functional brain mechanisms underlying effective intervention in children with MLD and provides novel metrics for assessing response to intervention.
The use of mathematics to categorize, visualize, and manipulate information extends to virtually all domains of human activity in the modern world. Yet, mathematical difficulties are widespread in school-age children, adolescents, and even college students. Individuals with MLD have specific difficulties with numerical and arithmetic problem solving, despite age-appropriate schooling and absence of impairments in other cognitive domains. Although the prevalence rate of MLD is comparable, or even higher, to that of reading disabilities, MLD has received considerably less attention from the research community. This is surprising because, to an even greater extent than reading skills, mathematical abilities in early childhood have been shown to impact later academic and professional achievement, socioeconomic well being, and health outcomes. It is therefore imperative to increase mathematical competence in all children, but most importantly in those falling at the lower end of the distribution.

School-based interventions designed to strengthen mathematical problem-solving skills have been shown to be effective in improving performance in some children with MLD but the neurobiological consequences of these programs are still unknown. Although lesion studies of individuals with math impairments have primarily focused on the specific role of the parietal cortex, recent neuroimaging research has begun to suggest that MLD involves aberrations in multiple functional systems. These include brain systems implicated in visual form judgement and symbol recognition, anchored in the ventral temporal–occipital cortex; quantity and magnitude processing, anchored in the intraparietal sulcus region of the parietal cortex; as well as attention and working memory functions, supported by a frontoparietal control network. However, the profile of differences in brain activation is currently an unresolved issue, as previous research has reported both over- and underactivation in MLD, relative to control groups. This lack of consensus is likely due to the limited number of studies, the use of different cutoff criteria for identifying MLD groups, inadequate matching on general cognitive measures such as intelligence quotient (IQ), reading and working memory, as well as the use of wide age ranges, and diverse experimental tasks and control conditions. Despite the lack of consistency in the direction of brain activations, previous studies have hinted that individuals with MLD show atypical functional responses in multiple brain areas, and not just parietal regions involved in quantity manipulation.

The extent to which effective behavioral intervention can alter aberrant activations in distributed brain systems is currently unknown. Specifically, it is not known whether successful behavioural interventions can effectively normalize activity in the same systems that show aberrant functional responses in MLD, or whether children with MLD recruit atypical neural resources to achieve the same level of performance as typically developing (TD) children. A comprehensive quantitative characterization of functional brain changes following effective intervention is required to address these questions and to elucidate the mechanisms by which poor math problem-solving skills can be remediated in children with MLD. This, in turn, can help identify neural factors mediating individual differences in response to intervention.

Importantly, as with other learning disabilities, response to interventions varies considerably across individuals. Analyses of individual differences have pointed to behavioural factors, such as severity of symptoms at the beginning of treatment (that is, severity of math disability), or domain-general cognitive abilities as potential mediators of success or failure in response to population-wise effective interventions. Here we investigate the possibility that poor response to intervention may be associated with weak functional neuroplasticity in children with MLD.

We combined cognitive assessments with event-related functional magnetic resonance imaging (fMRI) acquired during arithmetic problem solving and brain changes associated with an intensive 8-week 1:1 math tutoring program. We used a tutoring program that combines conceptual aspects of number knowledge and speeded practice on efficient counting strategies and systematic learning of number families. These components are designed to facilitate arithmetical fluency, and have previously been validated in school settings (Fig. 1a,b). We studied a well-characterized group of 7–9-year olds with MLD (Supplementary Tables 1 and 2), and a group of TD children matched on age, gender, IQ, reading and working memory abilities (Supplementary Table 1). First, we used voxel-wise univariate fMRI analysis to identify brain systems that are aberrant in children with MLD during numerical problem solving, and to characterize over or underactivation profiles of such aberrations, relative to TD children. Second, by combining univariate analyses with multivariate techniques of brain patterns classification, we directly test three competing hypotheses of functional brain changes in children with MLD: the (i) neural normalization hypothesis, which posits that atypical brain responses in children with MLD before tutoring become more similar to, and statistically indistinguishable from TD peers after tutoring; (ii) persistent neural aberration hypothesis, which posits that, even if performance deficits are successfully remediated, children with MLD will continue to show atypical responses in the same brain areas that they did before tutoring; and (iii) neural compensation hypothesis, which posits that children with MLD, after tutoring, would recruit additional and distinct (compensatory) brain systems compared with TD children. Finally, we developed novel machine-learning-based quantitative metrics to investigate whether a systematic relation exists between tutoring-induced neuroplasticity and behavioural gains in children with MLD.

We find that, consistent with the multi-component neurocognitive model of MLD, before tutoring, the profile of brain responses in children with MLD during arithmetic problem solving is characterized by prominent and widespread aberrations in multiple functional systems serving quantity, visuospatial, attention, working memory and cognitive control processes, that support successful numerical problem solving. Critically, 8 weeks of 1:1 tutoring results in marked changes in brain responses across all these systems. These widespread changes are characterized by normalization of responses, such that brain activation patterns in MLD children are no longer discriminable from their TD peers after tutoring. Crucially, multivariate pattern analyses provide novel evidence that the degree of such widespread tutoring-induced functional brain plasticity is uniquely associated with individual differences in response to intervention in children with MLD.

Results
To investigate behavioural and neurobiological consequences of 8 weeks of 1:1 math tutoring, we examined differences in performance and task-based fMRI activation levels during arithmetic problem solving, between and within groups, before and after tutoring (see Methods). Because tutoring could also alter performance and brain function in TD children, tutoring-induced changes in the MLD group were assessed with respect to both (a) pre-tutoring and (b) post-tutoring sessions in the TD control group.
Tutoring normalizes performance in children with MLD. We first examined changes in math performance following 8 weeks of tutoring and, as predicted by evidence from previous school-based studies\(^\text{17,19,40}\), we found that this type of 1:1 tutoring is effective in children with MLD, as assessed by significant improvements in accuracy on the arithmetic problem solving verification task performed inside the scanner (\(t_{(28)} = 3.323, P = 0.005\), Cohen’s \(d = 0.86\)) (Fig. 1c). Moreover, we found evidence for performance normalization in MLD: before tutoring, children with MLD were significantly less accurate than their TD peers (\(t_{(28)} = -2.318, P = 0.028\), Cohen’s \(d = 0.85\)), while their accuracy performance after tutoring did not differ from TD children at pre-tutoring (\(t_{(28)} = 0.471, P = 0.64\), Cohen’s \(d = 0.17\)), or post-tutoring (\(t_{(28)} = -0.598, P = 0.55\), Cohen’s \(d = 0.22\); Fig. 1c). The TD group did not show significant accuracy gains after tutoring (\(t_{(14)} = 1.469, P = 0.16\), Cohen’s \(d = 0.37\)). However, they showed significantly reduced reaction times after tutoring (\(t_{(14)} = -4.951, P = 0.0001\), Cohen’s \(d = 1.29\)), while this reduction was not evident in the MLD group (\(t_{(14)} = -1.188, P = 0.25\), Cohen’s \(d = 0.33\)). The groups did not differ on reaction times before (\(t_{(28)} = -0.694, P = 0.49\), Cohen’s \(d = 0.25\)) or after (\(t_{(28)} = 1.267, P = 0.22\), Cohen’s \(d = 0.46\)) tutoring.

Critically, these results were replicated in a separate arithmetic problem-solving task performed outside the scanner in which, instead of verifying addition equations, children were asked to verbally generate the answer to addition problems (Methods). Here again, performance differences that were evident between MLD and TD groups before tutoring (\(t_{(25)} = -2.631, P = 0.014\), Cohen’s \(d = 1.01\)), were entirely absent after tutoring (\(t_{(25)} = -1.141, P = 0.26\), Cohen’s \(d = 0.44\); \(t_{(25)} = 0.007, P = 0.99\), Cohen’s \(d = 0.01\) for TD’s performance before and after tutoring, respectively; Supplementary Fig. 1). These data demonstrate that tutoring-induced performance normalization cannot be explained by confounding factors related to being in the scanner the second time such as greater familiarity and/or diminished anxiety, and furthermore, it is independent of the response mode and the format in which the problems are presented.

Tutoring normalizes brain activity in children with MLD. Before tutoring, relative to TD children, children with MLD showed higher activation levels in multiple brain systems. Specifically, they showed higher activity in the bilateral prefrontal cortices including the dorsolateral and ventrolateral prefrontal cortices, the bilateral anterior insular cortices (Fig. 2a, Supplementary Table 3), the bilateral superior frontal gyri, and the right orbitofrontal cortex (Supplementary Table 3). Higher activation levels in the MLD group, compared to TD peers, were also seen in the inferior and superior parietal cortices encompassing the left intraparietal sulcus (Fig. 2a, Supplementary Table 3), the left supramarginal gyrus, and the right precuneus (Supplementary Table 3), as well as in the ventral temporal–occipital cortex including the bilateral fusiform gyri and the right lingual gyrus (Fig. 2a, Supplementary Table 3). Subcortical areas, including the bilateral cerebellum, right subcallosal cortex, left putamen and left anterior hippocampus and adjoining...
Figure 2 | Normalization of aberrant functional brain responses in children with MLD after 8 weeks of math tutoring. (a) Before tutoring, children with MLD (n = 15) showed significant differences in brain activation levels compared with TD children (n = 15). Significant group differences were evident in multiple cortical areas in the Prefrontal Cortex, including the bilateral Dorsolateral Prefrontal Cortices (DLPFC), and the left Ventrolateral Prefrontal Cortex (VLPFC), as well as the bilateral Anterior Insular Cortices (AIC); in the Parietal Cortex encompassing the left Intraparietal Sulcus (IPS); and in the Ventral Temporal–Occipital Cortex including the right Fusiform Gyrus (FG). (b) After 8 weeks of tutoring, functional brain responses in MLD children (n = 15) normalized to the levels of TD children (n = 15). Height threshold P < 0.01, extent threshold P < 0.01, significant by whole-brain voxel-wise independent samples t-test.

Table 4. Normalization of brain activity patterns in MLD. To further characterize the nature of tutoring-induced functional brain plasticity in children with MLD, we conducted additional within-group analyses contrasting brain activity levels before and after tutoring. Compared with the post-tutoring session, functional activation before tutoring was characterized by greater signal intensity levels in the same brain systems described above, encompassing prefrontal, parietal and ventral temporal–occipital cortices (Fig. 3a, Supplementary Table 7). Crucially, children with MLD did not show greater functional activation in any of these or other brain regions after tutoring (Fig. 3b).

Taken together, these results provide convergent evidence for the hypothesis that 1:1 math tutoring can normalize activity levels in children with MLD, to the level of their TD peers, in multiple brain systems important for numerical problem solving that are aberrant in children with MLD.

Normalizations of brain activity patterns in MLD. To further evaluate the neural normalization hypothesis, and contrast it with the persistent neural aberration hypothesis in a quantitatively rigorous manner, we used Support Vector Machine (SVM) classification techniques, to assess whether functional brain activity patterns in the brain systems identified above (Fig. 2a) could be used, over and above signal level intensity differences, to discriminate children with MLD from their TD peers either before or after tutoring (Fig. 4a). Brain activity patterns (t-maps) were first z-transformed (demeaned and scaled by variance), thus ensuring that group differences are independent of differences in task-related activity levels. The z-transformed activity maps were used as input features to a pattern-based SVM classifier, and Leave-One-Out Cross-Validation (LOOCV) procedures were used to assess discriminability of task-related activation patterns between the MLD and TD groups. We hypothesized that if, before tutoring, and in line with our univariate analysis, children with MLD engage brain areas differently, their task-related brain activity patterns would be highly discriminable from the TD group. Consistent with this hypothesis, multivariate pattern analyses revealed high (83.33%), and significant (P = 0.01), classification
Figure 3 | Tutoring-induced functional brain plasticity in children with MLD. (a) Before tutoring, compared with post-tutoring, children with MLD (n = 15) exhibited overactivation in several cortical areas in the Prefrontal Cortex, including the bilateral Dorsolateral and Ventrolateral Prefrontal Cortices (DLPFC and VLPFC), and the bilateral Anterior Insular Cortices (AIC), in the Parietal Cortex, encompassing the bilateral Intraparietal Sulci (IPS); and in the Ventral Temporal–Occipital Cortices, including the bilateral Fusiform Gyri (FG). (b) Post-tutoring, compared to pre-tutoring, no brain areas showed greater activation in children with MLD (n = 15). Height threshold P < 0.01, extent threshold P < 0.01, significant by whole-brain voxel-wise paired-samples t-test.

Functional brain plasticity predicts MLD performance gains. Finally, to assess individual differences of tutoring-induced functional brain plasticity in children with MLD, we computed a brain-based distance metric defined as the Brain Plasticity Index—BPI—(Fig. 4c). Specifically, BPI was calculated for each MLD child by computing a multivariate spatial correlation between pre- and post-tutoring patterns of brain activity, and subtracting it from 1 (Methods). Thus, high levels of BPI reflect greater tutoring-induced functional brain plasticity associated with tutoring. We then examined whether BPI was related to individual differences in performance improvement in children with MLD. Our analysis revealed a significant positive correlation (r = 0.526, P < 0.05), such that children with MLD who showed greater tutoring-induced functional brain plasticity, as indexed by BPI, exhibited larger performance gains with tutoring (Fig. 4d).

Behavioural measures do not predict MLD performance gains. None of the standardized measures (IQ, working memory or math) collected before tutoring were related to performance improvement in children with MLD (all P > 0.15) (Supplementary Fig. 8). Our results confirm that in contrast to these cognitive assessments, multivariate quantitative measures of functional change in brain activity (Fig. 4c), can provide a more sensitive metric for how well children with MLD respond to an intervention.

Discussion

Consistent with previous school-based studies, we found that 8 weeks of 1:1 math tutoring focused on strengthening conceptual and procedural knowledge can effectively improve arithmetic problem solving skills in primary-school children with MLD (Fig. 1c, Supplementary Fig. 1). Importantly, we demonstrate that, in parallel with performance normalization, 1:1 tutoring elicits extensive functional brain changes in children with MLD, normalizing their brain activity to the level of neurotypical peers. Prominent differences in brain activation between MLD and TD groups in prefrontal, parietal, ventral temporal–occipital cortices that were evident before tutoring, were entirely absent after tutoring (Fig. 2). Remarkably, tutoring resulted in significant reductions of widespread overactivation in multiple neurocognitive systems important for numerical problem solving (Fig. 3b,c,22–24,42, and machine-learning algorithms revealed that brain activity patterns in MLD learners were no longer discriminable from those of their peers after tutoring (Fig. 4b). Finally, children with MLD who displayed greater tutoring-induced functional brain plasticity also exhibited larger performance gains (Fig. 4d), highlighting, for the first time, the behavioural significance of widespread brain changes in response to intervention.

Although MLD was initially conceptualized as a disorder of a single brain region characterized by a localized deficit in the intraparietal sulcus,33,48,49 more recently, prominent neurocognitive models of MLD have posited that the disorder stems from more extensive functional aberrations in a distributed network of brain areas encompassing not only posterior parietal, but also prefrontal, as well as ventral temporal–occipital cortices that are known to serve multiple cognitive functions necessary for successful numerical problem solving.1–22, Consistent with this view, before tutoring, we found that, during arithmetic problem solving, children with MLD showed differential and widespread overactivation in multiple neurocognitive systems (Figs 2a and 3a), likely reflecting the need for greater neural resources during arithmetic problem solving, rather than an inability to activate task-relevant brain areas.32,33,50,51

Furthermore, effective intervention in MLD markedly altered aberrant neural responses by inducing functional changes in multiple brain systems (Fig. 2). These functional changes were characterized by reduction of activation following tutoring (Fig. 3b) and were much more widespread than frontoparietal cortex reductions of activation previously reported in brief (typically 5 days) arithmetic training studies in neurotypical adults.25, Neurodevelopmental studies of neurotypical children and adults34, and computer-based training studies in children with low math abilities35,36. This suggests that a comprehensive tutoring program, based on effective school interventions17,19,40 and designed to strengthen both number knowledge and arithmetic fluency7–19,39–41, can induce changes across distributed brain systems that encompass multiple stages of the information processing hierarchy necessary for successful numerical problem solving.6,22,42

Crucially, our finding of no differences in brain activation between MLD and TD children after tutoring (Fig. 2b) was
confirmed by more stringent Bayesian estimation procedures (Supplementary Figs 4 and 5) and, consistent with the neural normalization hypothesis demonstrates that this type of 1:1 tutoring can restore brain activity in children with MLD to the level of their TD peers. This is in stark contrast to the neural compensation hypothesis, which predicts that effective tutoring would be supported by compensatory mechanisms in other brain regions, and we found no evidence of greater activity in any brain area, after tutoring, in children with MLD, compared with their TD peers (Fig. 2b). Furthermore, univariate analysis found no evidence for the persistent neural aberration hypothesis either, as tutoring altered activity levels in all these systems (Fig. 3b), and differences in task-related brain activity levels. These results generally, the multivariate and quantitatively rigorous approaches also activity patterns in children with MLD. Strong multivariate evidence for the neural normalization hypothesis (Fig. 4b) highlights the possibility that this type of 1:1 tutoring program might prove beneficial for future learning in these children, without placing additional burden on effortful processing resources. More generally, the multivariate and quantitatively rigorous approaches used here to investigate learning-induced functional brain plasticity associated with tutoring offer powerful new metrics for detecting individual differences in brain activity patterns associated with interventions. Finally, leveraging our quantitative approach, and after characterizing the nature of tutoring-induced functional brain activity patterns during arithmetic problem solving were indistinguishable between the two groups. These results demonstrate that MLD and TD groups can be discriminated with high levels of accuracy prior, but not after tutoring, consistent with the neural normalization hypothesis. Critically, they are inconsistent with the persistent neural aberration hypothesis, which would have resulted in high discriminability between brain activity patterns in the MLD and TD groups at both the pre- and post-tutoring sessions. It is important to note that because the brain activity maps used as inputs to the classifier were normalized before classification analysis, discriminability between groups is independent of differences in task-related brain activity levels. These results further highlight the robustness of our findings and suggest that 8 weeks of tutoring not only normalizes brain activation levels, but also activity patterns in children with MLD. Strong multivariate evidence for the neural normalization hypothesis (Fig. 4b) highlights the possibility that this type of 1:1 tutoring program might prove beneficial for future learning in these children, without placing additional burden on effortful processing resources. More generally, the multivariate and quantitatively rigorous approaches used here to investigate learning-induced functional brain plasticity associated with tutoring offer powerful new metrics for detecting individual differences in brain activity patterns associated with interventions. Finally, leveraging our quantitative approach, and after characterizing the nature of tutoring-induced functional brain
plasticity in MLD (Figs 2 and 3), we extended our brain-based multivariate framework to assess individual differences in response to intervention in children with MLD, and assess whether tutoring-induced functional changes are behaviorally meaningful. Specifically, we computed a BPI for each MLD child based on the multivariate spatial correlation between pre- and post-tutoring patterns of brain activity. Remarkably, individual differences in performance gains in children with MLD were strongly correlated with their BPI measures: children who demonstrated greater tutoring-induced functional brain plasticity (that is, greater brain normalization) also exhibited larger performance gains with tutoring (Fig. 4d). In contrast, performance gains in children with MLD were not related to general cognitive abilities (for example, IQ; Supplementary Fig. 8). Furthermore, neuropsychological measures of math abilities before tutoring also did not predict performance gains (Supplementary Fig. 8), further highlighting the robustness and specificity of our brain-behavioural findings.

Our quantitative approach for measuring brain changes following intervention represents a significant advance over previous approaches that have primarily used neuroimaging data to investigate the anatomical loci and direction of change in activation. To our knowledge, this is the first use of a rigorous quantitative approach to assess distributed changes in brain activity pre- and post-intervention in children with MLD, which in turn is essential for more fully characterizing brain mechanisms underlying individual differences in response to intervention.

Further studies are needed to examine transfer to untrained and more complex math problems, and to assess the long-term stability of tutoring-induced functional brain plasticity in individuals with MLD, to make further assertions about the sustained effects of this intervention in the remediation of MLD. Future work should also contrast different types of tutoring programs and intervention groups, to further disentangle cognitive, motivational and psychological factors associated with remediation of learning disabilities. Finally, long-term follow-up studies should also clarify the extent to which the BPI metric developed here can serve as a predictive biomarker of long-term response to treatment56.

In conclusion, our study provides the first evidence that 8 weeks of 1:1 targeted cognitive tutoring can successfully remediate both poor math performance as well as aberrant brain responses in primary-school children with MLD. Children with MLD showed strong tutoring-induced functional neuroplasticity across distributed brain systems that support quantity, visuospatial, attention, working memory and cognitive control processes necessary for successful numerical problem solving. Critically, our findings demonstrate for the first time that quantitative measures of functional brain plasticity can significantly predict individual differences in response to treatment in children with MLD, further highlighting the unique potential of systems neuroscience-based approaches to the advancement of educational practice in remediating MLD. More generally, the quantitative framework developed here is likely to be useful for investigating brain plasticity associated with response to intervention in other forms of learning disabilities as well as in psychiatric, neurological, neurodevelopmental and neurodegenerative disorders56.

**Methods**

**Participants.** A total of 46 children in their third grade of schooling (ages 7.5–9.6 years) were recruited from multiple school districts in the San Francisco Bay Area. All participants were right-handed and without medical, neurological or psychiatric illness. Informed written consent was obtained from the legal guardian of the child and all study protocols were approved by the Stanford University Review Board. All participants were volunteers and were treated in accordance with the American Psychological Association Ethical Principles of Psychologists and Code of Conduct.

Two of 16 children were excluded from the study because they did not meet inclusion criteria for (i) in-scanner motion parameters (total frames interpolated < 20%) and adequate whole-brain coverage, (ii) in-scanner accuracy performance (> 50%) and (iii) neuropsychological scores (Supplementary Methods). Thus, the analyses presented in this study used a sample of thirty children, 15 of whom were characterized as having MLD, and 15 who were determined to be TD control children. This sample size was based on previously published studies of children with MLD57,29,30,35. Bayesian analyses confirmed that before tutoring, between-group differences in brain activation could be detected with a confidence interval of (Supplementary Figs 4 and 5). Additional details regarding participant selection, demographic as well as cognitive assessments, are presented in the Supplementary Methods.

**Overall study design.** Figure 1a illustrates the tutoring study design. Demographic, neuropsychological, cognitive and brain-imaging measures were acquired from each participant prior to tutoring. Before commencing 8 weeks of 1:1 math tutoring, participants underwent an MRI scan session, which included both a fMRI as well as a structural MRI acquisition protocol. The fMRI tasks consisted of an arithmetic verification task (Addition task), where the child had to verify the validity of addition equations (for example, 3 + 4 = 7), and a number identity verification task (Control task), where the child had to assess the validity of number identity expressions (for example, 7 = 7). After successful completion of the MRI scan session, children started the 8-week math tutoring program. Tutoring sessions occurred three times per week and were each approximately 40–50 min in duration. A second MRI scan session using the same protocols took place after the 8 weeks of tutoring.

**Neuropsychological assessments.** Each child participated in a neuropsychological assessment session in which they were tested on the Wechsler Abbreviated Scale of Intelligence (WASI)57, the Wechsler Individual Achievement Test (WIAT-II)58 and the Working Memory Test Battery for Children (WMTB-C)59. Full-scale IQ was determined using the WASI; academic achievement in reading and mathematics was assessed using the WIAT-II; working memory was assessed using the WMTB-C. These standardized measures were acquired before tutoring and were not readministered because of their limited validity when repeated within a year. The Scale for Early Mathematics Anxiety was also administered before and after tutoring (that is, greater brain normalization) also exhibited larger performance gains with tutoring (Fig. 4d). In contrast, performance gains in children with MLD occurred three times per week and were each approximately 40–50 min in duration. A second MRI scan session using the same protocols took place after the 8 weeks of tutoring.

**Tutoring sessions.** All children took part in an 8-week math tutoring program adapted from MathWise17,19,40. The tutoring combined conceptual instruction with speeded retrieval of arithmetical facts. The focus of the tutoring included strengthening of number knowledge (for example, cardinality) and relations within and between operations (for example, commutativity and inverse relation between addition and subtraction) that facilitate the use of sophisticated counting procedures and retrieval-based processes. The tutoring also incorporated a strategic practice component, that is important for building automaticity40, and decreasing load on cognitive resources (for example, working memory and non-verbal reasoning)58. This practice was designed to promote quick responding and use of efficient counting procedures to generate as many correct responses as possible, which in turn supports the formation and strengthening of representations in long-term memory45,46. Similarly to MathWise40, the tutoring involved a total of 15–20 h of tutoring. Differently from MathWise, for which tutoring occurred over the course of 15–16 weeks, the present tutoring was condensed to 8–9 weeks. Thus, the current tutoring had longer sessions—from 40 min to 50 min—to equate overall time on tutoring. Specifically, the present tutoring consisted of 22 lessons of increasing difficulty (details of each lesson and the tutoring material are described in the Supplementary Methods).

**Outside-the-scanner arithmetic production task.** On the day of the MRI session, before entering the scanner, children also performed an arithmetic production task (for example, 4 + 3 = 7). A total of 24 problems, involving random pairs of integers from 2 to 19, with sums ranging from 6 to 25 were presented to the child on a computer screen. The larger operand was equally likely to appear in the first or second position, as in the arithmetic verification task performed inside the scanner. The child was required to accurately solve each problem without the use of paper and pencil and to verbally state the answer out loud. For each problem, the experimenter recorded the child’s response verbatim on paper. Proportion of correctly solved problems was computed as the outcome measure of interest. Three children (one in the MLD group, and two in the TD group) did not complete this task due to time constraints. Parametric tests were used for this analysis (two-sided t-tests), as data were normally distributed in both samples (all K-S tests for normality P values were > 0.628 for the MLD group and > 0.320 for the TD group for both pre- and post-tutoring session).
Brain imaging. Functional MRI data acquisition. fMRI data were acquired using whole-brain imaging with a T2*-sensitive gradient echo spiral in/out pulse sequence at a Siemens 3 T MRI system. The following parameters were used: echo time (TE) = 30 ms, repetition time (TR) = 2 s, flip angle = 80°, field-of-view = 200 mm, 29 axial-oblique slices parallel to the AC-PC, dimensions 3.125 × 3.125 × 4 mm with 0.5-mm skip. To reduce blurring and signal loss from field inhomogeneity, an automated high order shimming method based on a spiral acquisition was used to improve the spatial resolution of the fMRI data. Subjects were placed in an MRI-lined cabinet to minimize head movements.

Functional MRI tasks. The arithmetic verification task—Addition task—took place inside the scanner during task-based fMRI acquisition. This task consisted of two runs of arithmetic problem solving during which the child had to verify the solution, addition equations (for example, 3 + 4 = 7). Problems were presented in a fast event-related fMRI design with 12 single-digit problems per run. In each run, problems were presented horizontally in green lettering on a black background. In half of the problems, the answers presented were correct (for example, 2 + 4 = 6); in the remaining half, the answers presented deviated from the correct solution by ± 1 or ± 2 (for example, 3 + 5 = 7). Arithmetic problems with 1 or 0 as operands were excluded. The larger operand was equally likely to appear in the first or second position. Each trial started with a fixation asterisk that lasted for 0.5 s. Then, the problem was presented for a maximum of 9.5 s, during which time the child could make the response. The participant used a response box, and if the answer was correct or not, after the response, the problem disappeared from the screen and a black screen appeared until the time window was filled to 9.5 s. A set of 12 non-arithmetic problems was also presented during each run and they constituted the control task. These problems consisted of number identity verifications (for example, 7 = 7) and were randomly interspersed with the arithmetic problems. Incorrect responses were followed by a correction (that is, the whole length of any given trial) with a canonical hemodynamic response function and a temporal derivative to account for voxel-wise latency differences in hemodynamic response. Voxel-wise contrast and t-statistic images were generated for each participant by contrasting arithmetic verification (Addition) versus number identity verification (Control) conditions and were averaged across the two runs. Both correct and incorrect trials were modeled in the GLM (four sub-conditions: Addition correct, Control correct, Addition incorrect, Control incorrect). The final voxel-wise contrast and t-statistic maps were generated on the first two sub-conditions only: Addition correct and Control correct.

At a group level, differences in brain activation between the MLD and TD groups were compared at both sessions (pre- and post-tutoring separately) using a t-test on contrast images of the Addition versus Control conditions. Differences in brain activation between pre- and post-tutoring sessions were compared within each group using a t-test on the same contrast images. In both analyses effects were modeled with brain areas in which the MLD and TD groups differed in activation at pre-tutoring (Fig. 2a, Supplementary Table 3). These maps were first z-transformed (demeaned and scaled by variance), thus ensuring that group differences independent of differences in task-related activity were removed. The z-transformed activity maps were used as input features to a pattern-based SVM classifier46 and LOOCV procedures were used to assess discriminability of task-related activation patterns between the MLD and TD groups. The classifier distinguished MLD from TD children by making a classification decision based on values of the linear combination of these features. We employed a widely used linear-classifier (SVM Classification)46 that was best suited for our purpose of classification-based models that are based on large number of features (brain based), but a small number of training samples (n = 30 in our case), which is typically the case of human brain neuroimaging studies. We used the Matlab package LIBSVM (http://www.csie.ntu.edu.tw/ cjlin/libsvm) to fit the classifier. LOOCV was used to measure the performance of the classifier in distinguishing MLD from TD children, based on patterns of brain activity. In LOOCV, data are divided into N folds (here, n = 30). A classifier is built using n − 1 folds, leaving out one sample. The left out sample is then classified using this classifier, and the Classification Accuracy (CA) is noted. The above procedure is repeated N times by leaving out one sample each time, and finally an average CA is computed.

Structural MRI data acquisition. High-resolution T1-weighted images were acquired in each child at both scan sessions (that is, pre- and post-tutoring), to facilitate anatomical co-registration of fMRI maps. A spoiled-gradient-recalled inversion recovery three-dimensional (3D) MRI sequence with the following parameters was used: TR = 300 ms, TE = 8.4 ms; flip angle = 15°; 22-cm field of view; 132 slices in coronal plane; 256 × 192 matrix; 2 NEX, acquired resolution = 1.5 × 0.9 × 1.1 mm.

Functional MRI data preprocessing. Data were analysed using SPM8 (http://www.fil.ion.ucl.ac.uk/spm). The first five volumes were not analysed and were used for slice timing equilibration. A 2-mm cubic field-of-view was used for the subtraction of the mean of the first six volumes separately for each slice during reconstruction using a magnetic field map acquired automatically by the pulse sequence at the beginning of the scan35. Images were realigned to correct for motion, corrected for errors in slice timing, co-registered to each individual’s structural T1 images, spatially transformed to standard stereotaxic space (based on the Montreal Neurologic Institute coordinate system), resampled every 2 mm using sinc interpolation, and smoothed with a 6 mm full-width half-maximum Gaussian kernel to decrease spatial noise prior to statistical analysis. For co-registration, the individual’s highest quality-rated (that is, either pre- or post-tutoring) structural MRI sequence was used.

Translational movement in millimeters (x,y,z), and rotational motion in degrees (pitch, roll, yaw) were calculated based on the SPm8 parameters for motion correction of the functional images of each subject. Mean scan-to-scan displacement of movement did not exceed 1 mm for all participants in either session (that is, pre- or post-tutoring). To correct for deviant volumes resulting from scans in motion, we used de-spiking procedures similar to those implemented in AFNI48. Volumes with movement exceeding 0.5 voxels (1.562 mm) or spikes in global signal exceeding 5% were interpolated using adjacent unaltered volumes. No > 12% of total volumes per run were repaired in either group. Critically, no differences between the groups on the total percent of volumes repaired or on the volume of each event including post-tutoring (P = 0.52), post-tutoring (P = 0.09), or between sessions in either group: MLD (P = 0.62), TD (P = 0.8). Translational movement parameters (x,y,z), and rotational movement parameters (roll, pitch and yaw) did not differ between the groups at either pre- or post-tutoring sessions (all P > 0.06). Within each group, movement parameters did not differ between post-tutoring sessions (all P > 0.09). Finally, mean scan-to-scan displacement also did not differ between the groups, or sessions (all P > 0.12).

Univariate analyses. General linear model. Task-related brain activation was identified using the GLM implemented in SPM8. At an individual level, brain activity related to task conditions was modeled using a boxcar function of 9.5 s (that is, the whole length of any given trial) with a canonical hemodynamic response function and a temporal derivative to account for voxel-wise latency differences in hemodynamic response. Voxel-wise contrast and t-statistic images were generated for each participant by contrasting arithmetic verification (Addition) versus number identity verification (Control) conditions and were averaged across the two runs. Both correct and incorrect trials were modeled in the GLM (four sub-conditions: Addition correct, Control correct, Addition incorrect, Control incorrect). The final voxel-wise contrast and t-statistic maps were generated on the first two sub-conditions only: Addition correct and Control correct.

Multivariate analyses. Multivariate classification analysis. We used multivariate classification analysis to assess whether functional brain activity patterns could be used to discriminate children with MLD from their TD peers either before or after tutoring (Fig. 4c). Brain activity (task-related activity) was defined in t-maps (that is, the whole-brain level). These maps were thresholded (based), but a small number of training samples (n = 30 in our case), which is typically the case of human brain neuroimaging studies. We used the Matlab package LIBSVM (http://www.csie.ntu.edu.tw/cjlin/libsvm) to fit the classifier. LOOCV was used to measure the performance of the classifier in distinguishing MLD from TD children, based on patterns of brain activity. In LOOCV, data are divided into N folds (here, n = 30). A classifier is built using n − 1 folds, leaving out one sample. The left out sample is then classified using this classifier, and the Classification Accuracy (CA) is noted. The above procedure is repeated N times by leaving out one sample each time, and finally an average CA is computed. Permutation tests (10,000 permutations of class labels) were conducted to arrive at P values associated with CA.

Multivariate distance analysis. We used a multivariate distance analysis to assess the behavioral significance of tutoring-induced functional brain plasticity (Fig. 4c). Specifically, we computed a brain-based distance metric defined as the BPI. BPI was calculated for each individual with MLD by computing a multivariate spatial correlation between pre-tutoring patterns of brain activity and post-tutoring patterns of brain activity. As in the previous section, these maps were masked with brain area in which the MLD and TD groups differed in activation at pre-tutoring (Fig. 2a, Supplementary Table 3), and were z-transformed (demeaned and scaled by variance), thus ensuring that effects were independent of differences in task-related activity levels. These values were then subtracted from 1, to facilitate easier conceptualization of the measure: the more plastic the brain, the higher the BPI. Parameter correlation analyses were then performed between the computed BPI values and performance gains (that is, calculated by subtracting pre-tutoring accuracy values from post-tutoring accuracy values) to investigate whether BPI was related to individual differences in performance improvement in children with MLD.

References
1. Butterworth, B., Varma, S. & Laurillard, D. Dyscalculia: from brain to education. Science 332, 1049–1053 (2011).
2. Nelson, W. & Reyna, V. Nutritional influences on the development of numerical competence for health decision making. Ann. Behav. Med. 33, 58–58 (2007).
3. Peters, E. et al. Numeracy and decision making. Psychol. Sci. 17, 407–413 (2006).
4. Reyna, V. F. & Brainerd, C. J. The importance of mathematics in human and animal judgment: numeracy, risk communication, and medical decision making. Learn. Individ. Differ. 17, 147–159 (2007).
5. Parsons, S. & Bynner, J. Does Numeracy Matter More? (NRDC, 2005).
6. Menon, V. in: Oxford Handbook of Numerical Cognition (eds Kadosh, R. C. & Dowker, A.) (Oxford University Press, 2015).
14. Butterworth, B. & Kovas, Y. Understanding neurocognitive developmental disorders can improve education for all. Science 340, 390–395 (2013).

15. Duncan, G. J. et al. School readiness and later achievement. Dev. Psychol. 43, 1428–1446 (2007).

16. Organization for Economic Co-operation and Development Trends Neurosci. Educ. 472 (2009).

17. Fuchs, L. S. Multiple components of developmental dyscalculia. J. Exp. Psychol. Gen. 142, 573–603 (2013).

18. Cortes, C. & Vapnik, V. in Machine Learning Vol. 20 273–297 (Kluwer Academic Publishers, 1995).

19. Suppek, K. et al. Neural predictors of individual differences in response to math tutoring in primary-grade school children. Proc. Natl Acad. Sci. USA 110, 8230–8235 (2013).

20. Isaacs, E. B., Edmonds, C. J., Lucas, A. & Gadian, D. G. Calculation difficulties in children of very low birthweight: a neural correlate. Brain 124, 1701–1707 (2001).

21. Kucian, K. & von Aster, M. Developmental dyscalculia. Cereb. Cortex 15, 1779–1790 (2005).

22. Davis, J. et al. Aberrant functional activation in school age children at-risk for developmental dyscalculia. Neuroimage 56, 993–1003 (2013).

23. Ischebeck, A. et al. How specifically do we learn? Imaging the learning of multiplication and subtraction. Neuroimage 30, 1365–1375 (2006).

24. Wechsler, D. in Handbook of Mathematical Cognition ed. Alibali, M. & Atkinson, R. (Academic Publishers, 1995).

25. Geary, D. C. & Hoard, M. K. in The High Cost of Developmental Dyslexia ed. Eden, G. F. (John Wiley & Sons Press, in press).

26. Ansari, D. Effects of development and enculturation on number representation in the brain. Nat. Rev. Neurosci. 9, 278–291 (2008).

27. Rosenberg-Lee, M. et al. Neural predictors of individual differences in response to math tutoring in primary-grade school children. Proc. Natl Acad. Sci. USA 110, 8230–8235 (2013).

28. Davis, N. in Developmental dyscalculia. Cereb. Cortex 15, 1779–1790 (2005).

29. Kaufmann, L. et al. Meta-analyses of developmental fMRI studies investigating typical and atypical trajectories of number processing and calculation. Dev. Neuropsychol. 36, 763–787 (2011).

30. Siegler, R. S. in A framework for remediating number combination deficits. J. Cogn. Neurosci. 22, 860–874 (2010).

31. Callaway, E. C. et al. Aberrant functional activation in school age children at-risk for developmental dyscalculia. Neuron 45, 257–269 (2005).

32. Kucian, K. et al. Impaired parietal magnitude processing in developmental dyscalculia. Neuroimage Clin. 2, 351–372 (2015).

33. Fuson, K. C. & Kwon, Y. Korean children's single-digit addition and subtraction: numbers structured by ten. J. Res. Math. Educ. 23, 148–165 (1992).

34. Siegler, R. S. & Shrager, J. in Origins of Cognitive Skills (ed. Sophian, C.) 229–293 (Lawrence Erlbaum Associates Publishers, 1984).

35. Kim, D. H., Adalsteinsson, E., Glover, G. H. & Spielman, D. M. Regularized higher-order in vivo shimming. Magnet. Reson. Med. 48, 715–722 (2002).

36. Kao, M. H., Mandal, A., Lazar, N. & Stufken, J. Multi-objective optimal experimental designs for event-related fMRI studies. Neuroimage 44, 849–856 (2009).

37. Glover, G. H. & Lai, S. Self-navigated spiral fMRI: Interleaved versus single-shot. Magnet. Reson. Med. 39, 361–368 (1998).

38. Geary, D. C. & Widaman, K. F. Numerical Cognition—on the convergence of componential and psychometric models. Intelligence 16, 47–80 (1992).

39. Fuson, K. C. & Kwon, Y. Korean children’s single-digit addition and subtraction: numbers structured by ten. J. Res. Math. Educ. 23, 148–165 (1992).

40. Kucian, K., Loenneker, T., Martin, E. & von Aster, M. Non-symbolic numerical processing in children and its relation to mathematics achievement. Front. Psychol. 3, 162. doi:10.3389/fpsyg.2012.00162 (2012).

41. Young, C. B., Wu, S. S. & Menon, V. The neurodevelopmental basis of math anxiety in second and third graders and its relation to mathematics achievement. Front. Psychol. 3, 162. doi:10.3389/fpsyg.2012.00162 (2012).

42. Geary, D. C. & Widaman, K. F. Numerical Cognition—on the convergence of componential and psychometric models. Intelligence 16, 47–80 (1992).

43. Fuson, K. C. & Kwon, Y. Korean children’s single-digit addition and subtraction: numbers structured by ten. J. Res. Math. Educ. 23, 148–165 (1992).

44. Siegler, R. S. & Shrager, J. in Origins of Cognitive Skills (ed. Sophian, C.) 229–293 (Lawrence Erlbaum Associates Publishers, 1984).

45. Kim, D. H., Adalsteinsson, E., Glover, G. H. & Spielman, D. M. Regularized higher-order in vivo shimming. Magnet. Reson. Med. 48, 715–722 (2002).

46. Kao, M. H., Mandal, A., Lazar, N. & Stufken, J. Multi-objective optimal experimental designs for event-related fMRI studies. Neuroimage 44, 849–856 (2009).

47. Glover, G. H. & Lai, S. Self-navigated spiral fMRI: Interleaved versus single-shot. Magnet. Reson. Med. 39, 361–368 (1998).

48. Cox, R. W. AFNI: software for analysis and visualization of functional magnetic resonance neuroimages. Comput. Biomed. Res. 29, 162–173 (1996).

Acknowledgements

This work was supported by NIH grants (HD047520, HD059205, HD057610 to VM; MH01094 to MRI; and HD053714 to LF). We thank Drs Daniel Abrams, Christian Battista, and Tanya Evans for useful feedback.

© 2015 Macmillan Publishers Limited. All rights reserved.
Author contributions
T.I., M.R.L., C.T., I.F. and V.M. designed the research; T.I., J.R., C.T. performed the research; T.I., M.R.L. and K.S. analysed the data; and T.I. and V.M. wrote the paper.

Additional information
Supplementary Information accompanies this paper at http://www.nature.com/naturecommunications

Competing financial interests: The authors declare no competing financial interests.

Reprints and permission information is available online at http://npg.nature.com/reprintsandpermissions/

How to cite this article: Iuculano, T. et al. Cognitive tutoring induces widespread neuroplasticity and remediates brain function in children with mathematical learning disabilities. Nat. Commun. 6:8453 doi: 10.1038/ncomms9453 (2015).

This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/