The spatiotemporal characteristics of multi-demand brain networks and their relationship with mathematical learning

Tongxi Yang, Yixuan Sun, Rui Guo

December 9, 2024

Introduction

As one of the basic subjects of modern education, mathematics is not only of great significance to social development, but also the cornerstone of modern science. It also has extraordinary significance for the survival and development of individuals. Studies have shown that even if reading ability, individual intelligence and economic factors are excluded, arithmetic and basic geometry ability will still affect individual employment, wages and work output (Rivera-Batiz, 1992; Rourke & Conway, 1997). However, in mathematics education, we have seen huge individual differences. There are both gifted children with excellent mathematics performance and children with difficulties in mathematics learning. They even have difficulty solving the simplest arithmetic problems (Butterworth, 2010; Iuculano et al., 2008; Kaufmann et al., 2013). Therefore, it is necessary for us to study the cognitive factors that affect mathematics learning and their corresponding neural mechanisms. This will not only help enhance the self-worth of backward children, enable them to contribute their intelligence and talents to society, but also help promote the rapid development of science.

1.1 Cognitive mechanism of mathematics learning

1.1.1 Mathematics learning has the characteristics of multi-step and flexibility and involves multiple cognitive and emotional factors

Mathematics learning requires the participation of multiple cognitive and emotional factors. Mathematics learning is a complex cognitive process that requires multiple cognitive components participation, such as working memory, attention, emotion and other cognitive factors. In the past few decades, cognitive neuroscience research on mathematics has mostly focused on the impact of a single cognitive component on mathematics academic performance, hoping to identify cognitive factors that have a greater impact on mathematics academic performance (Floyd et al., 2003; Geary, 2011; Kyttälä & Lehto, 2008). Studies have shown that working memory is related to the conceptual understanding and problem-solving process of mathematical learning. The degree of working memory participation required is positively related to the difficulty of problem solving. At the same time, individual differences in working memory are significantly related to differences in mathematical ability (Menon , 2016; Swanson, 2011).Attention also has a significant impact on academic performance in mathematics, and this effect persists even after controlling for working memory (Swanson, 2011).In addition, mathematics learning is also related to individual risk-taking behavior. Individuals with risk-taking tendencies will be more proactive in challenging more difficult content in the process of mathematics learning and will also be more persistent in solving difficult problems, and thus have higher mathematics achievements (Erbas & Bas,2015). Generally speaking, mathematics learning places higher demands on individual cognitive abilities.

1.1.2 Developmental effects of mathematics learning

Elementary school is not only the basic stage for individuals to learn mathematics, but also a stage for the rapid development of various cognitive and emotional factors of individuals. Many basic skills in mathematics are formally taught at the primary school level, and these skills are the basic framework of the entire mathematics edifice and have an important impact on individuals' lifelong learning. For example, arithmetic skills, which are only formally learned in primary school, are significantly related to academic performance in mathematics, and can predict subsequent individual career development. In addition, the primary school stage is also a stage of rapid development of individual cognitive and emotional abilities, so it is necessary to explore the impact of cognitive and emotional development on mathematics learning (Hao et al., 2021).

1.2 Neural Mechanisms of Mathematical Learning

1.2.1 Mathematical Learning Involves a Wide Range of Cerebral Cortex

The intraparietal sulcus is a unique brain area for mathematical learning. Quantity is the most basic content in mathematical cognition and the basis for the development of many subfields of mathematics. Many animals, like humans, have the ability to represent quantity approximately and have a distance effect. When performing tasks such as quantity comparison, they will activate the intraparietal sulcus-related areas (Cantlon, 2012). In addition, humans will also activate the intraparietal sulcus-related areas when performing numerical comparison tasks. A recent meta-analysis on mathematical cognition showed that when people perform arithmetic and spatial cognition, although the activation patterns are quite different, they will activate the intraparietal sulcus area (Hawes et al., 2019). This may indicate that the intraparietal sulcus is a functional brain area that is unique to mathematical cognition and is the basis of mathematical learning.

1.2.2 Mathematical learning and cerebellum

The cerebellum is usually considered to be the neural basis related to motor control, but with the development of neuroimaging technology in recent years, people have gradually discovered that the cerebellum plays an important role in information processing. A recent machine learning article on cortex-cerebellum coupling showed that the feedback mechanism of the cerebellum to the brain helps individuals learn sensory motor and cognitive tasks (Boven et al., 2023). In addition, neuroanatomical studies have also discovered the physiological mechanism of this feedback. The cerebellum often mediates information processing by coupling with the cerebral cortex. Fibers from the cerebral cortex are projected to the cerebellum through the pons; and the cerebellum returns to the cerebral cortex through the thalamus, eventually forming a loop (Bellebaum &

Daum, 2007). This gives us a new understanding of the computational mechanism of the cerebellum.

1.3 Neural Mechanism of Multi-demand System

1.3.1 Cognitive Mechanism of Multi-demand System

Human goal-oriented behavior is often achieved by completing a series of sub-goals. For example, when people solve mathematical problems, they must first reasonably represent the problem and list the correct formula. Then they need to calculate according to the operation rules. Finally, they need to connect the calculation results with the actual problem and make a correct explanation. To achieve this goal, people must focus on achieving each sub-goal in chronological order, and then pass the results of each stage to the next sub-goal. In addition, when people learn a new task, they need to break down and define the task into a series of useful sub-tasks. This ability is the basis of many complex human behaviors (Duncan, 2010, 2013).

1.3.2 The relationship between mathematics and multi-demand brain networks

Mathematics is a high-level cognition unique to humans. Whether in the process of learning mathematics or applying mathematics, it requires the participation of multiple cognitive components and is accompanied by a high level of cognitive control. In addition, research shows that an individual's fluid intelligence level is significantly related to mathematics achievement (Desco et al., 2011). Therefore, we believe that the cognitive mechanism of the multi-demand brain network is very consistent with the cognitive needs of mathematics learning and application and may be an important neural mechanism that promotes mathematics learning. However, previous studies on multi-demand brain networks have focused on speech understanding (Blank et al., 2014; Diachek et al., 2020). Research by Evgeniia (2020) shows that multi-demand brain networks are not involved in the core part of speech understanding. This reminds us of Almaric's series of studies on brain networks specific to mathematical knowledge (Amalric & Dehaene, 2016, 2018, 2019). This network has great overlap with MDS.

In summary, researchers focus more on the impact of a single cognitive factor on mathematics learning, such as the impact of working memory on mathematics academic achievement. However, mathematics learning is a high-level human cognitive activity, which is affected by multiple cognitive factors. Only by considering multiple cognitive factors together can we better understand the essence of mathematics learning, and the multi-demand brain network is the common neural basis of multiple cognitive factors. We will explore the impact of the multi-demand brain network on mathematics learning.

Data introduction

2.1 MRI Task

The data for this study is from the Beijing section of the Children's School Functions and Brain Development Project (CBD). The database recruits children of appropriate age in primary and secondary schools in Beijing, and collects behavioral and imaging data from three units: Peking University, Beijing Normal University, and Huilongguan Hospital.

After matching the questionnaire and behavioral data, the analysis included brain imaging and behavioral data of 310 school-age children (6-12 years old), including 147 boys and 163 girls. All subjects had no visual impairment after screening and recalibration and had no history of any neurological or psychiatric diseases, and no medication experience during the study. The child subjects and their parents have been informed of the experimental process and signed the informed consent form in writing. The MRI task-state data collection includes the following four tasks.

2.1.1 Working memory task

Working memory is assessed using the N-back task, which measures an individual's working memory ability by controlling different working memory load conditions. The task consists of three different workload conditions (0-back, 1-back, and 2-back), each condition consisting of 4 blocks, for a total of 12 blocks. In each block, the subject first sees a prompt about the current block (0-back, 1-back, and 2-back) for 2 seconds. Then a sequence of 15 random numbers is presented, each of which is displayed for 400 milliseconds. In the case of 0-back, the subject needs to respond to the number "1" by pressing the key. In the condition of 1-back, the subject needs to respond to the number that is the same as the previous number. In the condition of 2-back, the subject needs to respond to the number that is the same as the two previous numbers.

2.1.2 Emotion recognition task

The experimental paradigm of emotion recognition uses the Emotion Face Matching task, which consists of an emotion condition and a control condition. Each condition contains 5 blocks, and each block contains 6 trials. In each block, the subject first sees a prompt about the current block (emotion matching/graphic matching), indicating whether the next task is to identify facial emotions or graphic directions. The prompt is presented for 5 seconds. In each block, 6 trials are presented one by one, each trial is presented for 5 seconds. The subject needs to choose the stimulus with the same emotion or graphic direction as the target stimulus from the two stimuli.

2.1.3 Attention ability task

The experimental paradigm of attention ability adopts the Attention Network Test, which consists of a cue presentation phase (no cue, double cue, spatial position cue and center cue) and a flanker task phase (consistent and inconsistent), with a total of six different trials. Because the task lasts for a long time, it is divided into two phases, each including 96 trials. At the beginning of each trial, the subject first sees a fixation point, which lasts between 400 and 1000 milliseconds. Then the cue will be presented for 150 milliseconds. The clues are divided into: (1) no cue; (2) double cue: two asterisks appear in the upper and lower positions; 3) spatial cue: one asterisk appears in the upper or lower position; (4) center cue: one asterisk appears in the center position. Then a fixation point is presented again for 450 milliseconds. Next, the subject will see a row of small fish and needs to judge the direction of the fish in the middle position. The directions of the surrounding non-target small fish and the target small fish may be consistent or inconsistent, which constitutes the two conditions of the flanker condition. Feedback will be presented after 999 milliseconds. If the subject responds correctly, the fish in the middle position will spit bubbles.

2.1.4 Risk decision-making task

The experimental paradigm of risk decision-making uses the Balloon Analogue Risk Task, which simulates the decision-making process of risk situations in real life by inflating a balloon. In each trial, the subject needs to press a button to inflate the balloon (key 1) or stop inflating (key 2). Each time the balloon is inflated, it expands a little bit, and the subject's reward increases by 1 yuan. The larger the balloon, the greater the possibility of explosion. If it exceeds the preset maximum value, it will explode. If the subject does not respond for a long time (3 seconds), the balloon will also explode, and the reward obtained in this round will be cleared. The subject needs to obtain the maximum benefit by inflating and stopping inflating at the appropriate time. The stage of choosing whether to inflate is the "reward-risk trade-off" condition, the interface of reward feedback that appears after the subject decides to stop inflating is the "reward processing" condition, and the interface of balloon explosion is the "risk processing" condition.

2.2 Academic Achievement Test

The content of the mathematics academic achievement test includes three aspects: numbers and algebra, space and graphics, statistics and probability. It examines students' four abilities: knowing facts, applying rules, mathematical reasoning, and solving unconventional problems. In order to avoid the "floor effect" or "ceiling effect", this test is divided into two stages according to the "Mathematics Curriculum Standards". Different test books are used for different stages. Grades 1-3 are the first stage, and grades 4-6 are the second stage. Mathematics scores are standardized within the stage, and the final CEEB score is 500 with a standard deviation of 100.

Method

Figure: overall workflow

3.1 Preprocessing

All MRI data were processed using the same standardized preprocessing process fMRIPrep1.4.1 (Esteban et al., 2019). For the sake of subject adaptation and signal stability, the first four images of all task images were removed. The image sequence was then time-aligned and the image deformation was adjusted according to the field map. Finally, the task image was registered with the structural image and further standardized to the standard spatial template (Montreal Neurological Institute, MNI). After smoothing with a Gaussian kernel with a full-width half-maximum (FWHM) of 6 mm, the noise was further processed using ICA-based Automatic Removal of Motion Artifacts (ICA-AROMA), and the average signal of white matter and cerebrospinal fluid and various head motion parameters were output. Images with a displacement greater than 0.5 mm or a whole-brain standard signal change greater than 1.5 were marked as abnormal.

3.2 ROIs extraction and FC calculation

Before formally analyzing the static functional connectivity of the brain, the average signals of white matter and cerebrospinal fluid in the preprocessing output results, and

24 columns of head motion regression were regressed. Next, the Shen 268 template was used to extract the average BOLD signal of all voxels in the regions of interest (ROI), then we map 268 ROIs to 8 networks, and the Pearson correlation between brain regions was calculated. After removing the diagonal elements of the matrix, the lower triangular elements of the matrix were used for prediction analysis, and each subject had 35778 features.

3.3 Model

3.3.1 CPM

Connectome-based predictive modeling (CPM) is a data-driven framework designed to develop predictive models that elucidate brain–behavior relationships using connectivity data. This protocol comprises four key steps: (i) feature selection, (ii) feature summarization, (iii) model construction, and (iv) evaluation of prediction significance. Empirical evidence has demonstrated that CPM performs on par with, or even surpasses, other established approaches in brain–behavior prediction tasks. The specific process is shown in the figure below. For a more detailed introduction, please refer to the original protocol.
Figure: The framework of CMP
Considering the characteristics of our data, we noted a high-dimensional feature space comprising approximately 35,000 features and a relatively small sample size of 300 subjects. To solve this issue, we implemented dimensionality reduction by employing partial least squares (PLS) regression in place of ordinary least squares (OLS). Additionally, based on prior evidence and our own observations, we hypothesized that positive and negative edges within the connectivity matrix might have distinct effects on model performance. To account for this, we summed positive and negative edges separately during feature summarization. And we incorporated age and gender as covariates to control for their confounding effects. To avoid overfitting, we used 10-fold cross validation to fit the prediction model. For computational efficiency considerations, this study only analyzed the distribution pattern of the first principal component. We sorted the load coefficients of the functional connections of the four tasks with the first 10% as the threshold, screened out the edges that were important to the four tasks, and performed weight analysis on the nodes involved to obtain a series of important nodes.
Figure: Framework for validating the effectiveness of four tasks
Figure: Fig. x. Framework of non-linear model method. A Data Integration and Feature Engineering Workflow. B Optimal Model Structure and Hyperparameters Identification. C Feature Mask Layer and Feature Importance Assessment. D Network Localization of Top ROIs.

3.3.2 Data Integration and Feature Engineering for Enhanced Model Performance

In order to enhance the diversity of the data set and to enable the model to extract unique features from multiple datasets, we elected to merge four task-based datasets. The objective of this consolidation was to provide a more comprehensive view of the data, thereby enabling the model to capture a richer feature set. To mitigate feature interference, we divided the functional connectivity (FC) into two distinct subsets: "inner sub-feature sets," representing the FC within each network's regions of interest (ROIs), and "outer sub-feature sets," representing the FC between different network pairs' ROIs.

The aforementioned method of splitting the dataset serves to reduce interference between regions of interest (ROIs) while simultaneously preserving the connectivity features between individual brain networks (Fig. xA; Zhu et al., 2021).
To more effectively examine the influence of ROIs within diverse networks and their interconnections on network performance, we utilized the Partial Least Squares (PLS) model to conduct independent training on each subset of features. The loadings corresponding to each ROI were extracted from the model and applied as weights to the features. In order to identify the optimal parameters for each PLS training, a grid search was employed, specifically to determine the number of principal components to be extracted. Furthermore, 5-fold cross-validation was utilized to mitigate the risk of over-fitting (Fig. xA).Furthermore, an attempt was made to construct a stacking model comprising an elastic-net regression model as the meta-model and PLS as the base model. However, this approach yielded unsatisfactory results, indicating that the internal and external connectivity features of the network were not effectively leveraged.

3.3.3 Graph Attention Network (GAT) model

To address the limitations of the Partial Least Squares (PLS) model, which is adept at capturing only linear correlations within features, we introduced a nonlinear model to detect nonlinear patterns within the functional connectivity matrices. This strategic choice was essential to account for the complex, non-linear interactions inherent in the data. In light of these considerations, the Graph Attention Network (GAT) model was selected for its capacity to integrate both internal and external connectivity features and to capture non-linear and deep-level feature information (Velickovic et al., 2017). The methodology entailed treating each network as a node within the GAT framework, with the inner sub-features serving as node attributes. The interconnections between networks were regarded as edges, with the outer sub-features functioning as edge attributes. Following a comprehensive comparison of various model structures and a meticulous search for optimal grid parameters (see Fig. xB and Fig. xC), we have identified the most effective model structure and hyperparameters, as illustrated in the Fig. xB.
The Feature Mask layer is employed to ascertain the relative importance of input features. Throughout the training process, these parameters are updated in accordance with the gradients of the loss function, thereby enabling the model to discern more pertinent features with values approaching 1 from those of lesser significance with values approaching 0 (Ding et al., 2020). The deployment of Feature Masks in conjunction with regularization techniques prompts the model to drive unimportant feature mask values towards 0, thereby facilitating the generation of a sparse and focused representation of the most relevant features. We then extracted the top 10% of important features using the Feature Mask and identified the networks where these features are predominantly located.

3.3.4 t-test

To explore the individual differences in multi-demand networks, in our study, we employed quartile analysis to categorize participants. Similarly, their fc_matrix is converted into networks, and a t-test is performed on the high and low groups. As shown in the figure, specifically, in the attention task, the differences in Medial frontal-Medial frontal and Frontoparietal-Medial frontal are significant, which are located in the multi-demand networks. In other tasks, we didn't find differences in multi-demand networks.

Figure: t-test fig

3.3.5 Differences across the brain

In order to explore whether there are significant differences in overall brain connectivity between high and low groups, we refer to the paper below and choose connection density as an indicator, which can represent the information transmission and functional integration capabilities of the brain. As shown in the table, we couldn't identify differences in connectivity across the whole brain, but we could in the multi-demand networks, which means the multi-demand networks are associated with mathematical learning.

Table: t-test

Results

Study 1: The relationship between multiple tasks and mathematics performance.

4.1 The relationship between FC of the four tasks and mathematics performance

First, we examined the relationship between the four task FCs and mathematics performance. The results are shown in Table.

4.1.1 N-back task

Table: N-back task

4.1.2 Attention Network Test

Table: Attention Network Test

4.1.3 Emotion face matching task

Table: Balloon Analogue Risk Task

4.1.4 Emotion face matching task

Table: Emotion face matching task

4.1.5 summary of four tasks separately

Table: summary of four tasks separately

4.2 Enhanced Performance of GAT Model in Mathematics-Related Brain Network Analysis

In stark comparison to the linear Partial Least Squares (PLS) model, the Graph Attention Network (GAT) model has demonstrated superior performance in the task at hand. A comprehensive comparison of the specific performance metrics is delineated in Table X. The GAT model, when tasked with aggregation, has achieved performance metrics of R-squared = 0.6259 and Mean Squared Error (MSE) = 4461.43, which signifies a notable

enhancement over the performance metrics of models Lm1 and Lm2.
Table: Comparison of model performance metrics on different tasks.
Upon mapping the extracted features to their respective networks, it was ascertained that the networks implicated in mathematical academic performance are predominantly localized within the multi-demand system, particularly in the cerebellum, and to a lesser extent, in the connectivity between the cerebellum and Visual II region (as illustrated in Figure xD). This outcome aligns with our prior research. Furthermore, our findings suggest that a holistic assessment of all tasks may yield more profound insights compared to a singular focus on a specific task.

Study 2: Differences in the multi-demand networks

4.3 YiXuan

The conclusion is that we found that the neural networks related to math learning performance are mainly located in the multiple demand system. It is worth to note that in attention task, the functional connectivity of the multiple demand network showed significant variability. Based on these findings, we suggest that students can use these objective assessments to understand their own ability characteristics, then they can combine their personal interests, and choose the development path that best suits them.

Discussion

5.1 Summary

Improving mathematics performance involves more than repetitive practice. Our study highlights that integrating multiple tasks provides a more comprehensive approach to identifying the factors influencing math performance compared to relying on a single task. By completing four distinct tasks, it is possible to pinpoint specific cognitive deficiencies that may contribute to poor math outcomes. These findings suggest that educational interventions should focus on targeted training programs designed to address the specific deficits identified through task performance, offering a more personalized and effective strategy for improving mathematical abilities.
Brain networks appear to differ between individuals with high and low mathematical performance, particularly within the multi-demand system. Our studies reveal that functional connectivity in the multi-demand networks, especially during attention tasks, shows significant variability between these groups. This suggests that the multi-demand system plays a critical role in math-related abilities. To support students, promoting self-awareness of their cognitive strengths and weaknesses can be beneficial. By understanding their abilities and personal interests, students can make informed decisions about their developmental pathways, optimizing both academic and personal growth.

5.2 Highlight and limitation

Our study identifies a specific set of functional connections within the multi-demand network that precisely represents mathematical cognition. The multi-demand network highlights the relationship between various cognitive functions—such as attention, emotion, working memory, and risk decision-making—and mathematical learning. Furthermore, our research provides practical recommendations for improving mathematics education

and learning strategies. By adapting the model framework to the characteristics of our data, we developed a model that is both interpretable and effective.
Currently, our analysis of brain activity patterns is restricted to static functional connectivity. Since the brain operates as a dynamic system, a dynamic perspective would provide a more accurate understanding of the working mechanisms of the multi-demand network. Additionally, our exploration of brain network differences between high- and low-performing individuals remains preliminary. Future studies could incorporate graph theory to conduct a more comprehensive analysis of these network differences, potentially uncovering deeper insights into the neural basis of mathematical cognition.

Contributions of each member

Table: contributions of each member

References