Cerebral Cortex, 2023, **33**, 881–894
https://doi.org/10.1093/cercor/bhac108
Advance access publication date 7 March 2022
Original Article

A connectome-based neuromarker of nonverbal number acuity and arithmetic skills

Dai Zhang[1,2], Liqin Zhou[1], Anmin Yang[1], Shanshan Li[1], Chunqi Chang[2], Jia Liu[3], Ke Zhou[1,*]

[1]Beijing Key Laboratory of Applied Experimental Psychology, School of Psychology, Beijing Normal University, No. 19, Xinjiekouwai Street, Haidian District, Beijing 100875, China,
[2]Guangdong Key Laboratory for Biomedical Measurements and Ultrasound Imaging, School of Biomedical Engineering, Shenzhen University Health Science Center, No. 1066, Xueyuan Street, Nanshan District, Shenzhen 518060, China,
[3]Department of Psychology & Tsinghua Laboratory of Brain and Intelligence, Tsinghua University, No. 30, Shuangqing Street, Haidian District, Beijing 100084, China
*Corresponding author: Beijing Key Laboratory of Applied Experimental Psychology, School of Psychology, Beijing Normal University, No. 19, Xinjiekouwai Street, Haidian District, Beijing 100875, China. Email: kzhou@bnu.edu.cn

The approximate number system (ANS) is vital for survival and reproduction in animals and is crucial for constructing abstract mathematical abilities in humans. Most previous neuroimaging studies focused on identifying discrete brain regions responsible for the ANS and characterizing their functions in numerosity perception. However, a neuromarker to characterize an individual's ANS acuity is lacking, especially one based on whole-brain functional connectivity (FC). Here, based on the resting-state functional magnetic resonance imaging (rs-fMRI) data obtained from a large sample, we identified a distributed brain network (i.e. a numerosity network) using a connectome-based predictive modeling (CPM) analysis. The summed FC strength within the numerosity network reliably predicted individual differences in ANS acuity regarding behavior, as measured using a nonsymbolic number-comparison task. Furthermore, in an independent dataset of the Human Connectome Project (HCP), we found that the summed FC strength within the numerosity network also specifically predicted individual differences in arithmetic skills, but not domain-general cognitive abilities. Therefore, our findings revealed that the identified numerosity network could serve as an applicable neuroimaging-based biomarker of nonverbal number acuity and arithmetic skills.

Key words: connectome-based predictive modeling; approximate number system; arithmetic skills; individual differences; resting-state fMRI.

Introduction

The number of items in a set is represented in an approximate number system (ANS) shared by adults (Barth et al. 2003; Feigenson et al. 2004; Pica et al. 2004), infants (Xu and Spelke 2000), many animal species (Nieder and Miller 2004; Piffer et al. 2012; Agrillo et al. 2016), and even deep neural networks (Nasr et al. 2019; Kim et al. 2021; Zhou et al. 2021), which encodes quantity in an approximate, nonsymbolic manner (i.e. numerosity). The ANS is of evolutionary importance for humans and animal species (Gross et al. 2009; Geary and Moore 2016). Particularly, in humans, although the association between nonverbal number acuity and symbolic math performance remains controversial (Brankaer et al. 2013; Libertus et al. 2013; Bull and Lee 2014; Chen and Li 2014; Fazio et al. 2014; Leibovich-Raveh et al. 2017; Schneider et al. 2017), especially in adults, there is some evidence of a close link between nonverbal ANS acuity and math achievement in children (Halberda et al. 2008), suggesting that the ANS may play a bootstrapping role in the development of mathematical abilities (Piazza 2010).

Over the last decades, considerable progress has been made in the understanding of how numerosity information is represented in the brain. For instance, both neurophysiological and human functional magnetic resonance imaging (fMRI) studies have revealed that neurons or regions in the prefrontal and parietal cortices, including the horizontal part of the intraparietal sulcus, are tuned to the numerosity of spatial arrays (Piazza et al. 2004; Nieder 2012; Viswanathan and Nieder 2013; He et al. 2015; Nieder 2016; Harvey et al. 2017; Nieder 2017; Hawes et al. 2019; Ramirez-Cardenas and Nieder 2019). However, there are also new evidence showing that processing of numerosity information involves other brain areas (Dehaene and Cohen 1997; Dehaene et al. 2003; Park et al. 2015; Harvey 2016; Fornaciai et al. 2017; Kutter et al. 2018; Lasne et al. 2019). For example, a numerical neural system proposed by Harvey et al. (2016) consists of the visual cortex, inferior temporal gyrus (ITG), fusiform gyrus, angular gyrus, parietal lobe, and prefrontal cortex. It thus suggests that the neural representation of the ANS system may involve multiple cortical areas. However, most of these previous studies focused

Received: December 19, 2021. **Revised:** February 15, 2022. **Accepted:** February 16, 2022

on identifying the function of specific brain regions supporting the ANS, and a connectome-wise neuromarker based on whole-brain functional connectivity (FC) that can characterize an individual's number acuity is still lacking.

The FC measures the synchronization of fMRI signals between different brain areas (Biswal et al. 1995; Goense et al. 2012), providing information on how distributed brain regions integrate to support specific cognitive functions (Smith et al. 2013). Recently, a connectome-based predictive modeling (CPM) analysis (Shen et al. 2017) was proposed to predict individual differences in human behavior (e.g. cognitive abilities) based on the whole-brain FC pattern of participants. The CPM analysis provides a complementary measure of human cognitive abilities to traditional approaches that focus on the specific functions of discrete brain regions and has been successfully applied to predict individual differences in several cognitive abilities (Finn et al. 2015; Rosenberg et al. 2016; Beaty et al. 2018; Jangraw et al. 2018). Because substantial variations also exist in ANS acuity among individuals (Halberda et al. 2008; Lukowski et al. 2017; Wang, Halberda, et al. 2017a), here we aimed to employ the CPM analysis to the resting-state fMRI data to develop a whole-brain neuromarker (i.e. numerosity network) of ANS acuity.

In addition, previous studies also illustrated that the CPM identified using one task could be generalized to predict the performance of other related tasks, indicating a shared neural basis between these tasks (Rosenberg et al. 2016; Jangraw et al. 2018). Considering the above-mentioned relationship between ANS acuity and mathematical abilities, our second aim was to examine whether the identified numerosity network could predict an individual's arithmetic skills, and, if so, whether the predictive ability of the numerosity network is specific to arithmetic skills or is general for broader domain-general cognitive abilities, such as fluid intelligence (gF), nonverbal working memory, executive functioning, or language comprehension.

Specifically, in a large sample dataset ($n > 250$), we measured each individual's ANS acuity using a dot-array number-comparison (NC) task (Halberda et al. 2008). We then constructed a numerosity CPM based on their resting-state fMRI data to predict individual differences in ANS acuity using a leave-one-out (LOO) cross-validation procedure. We found that the summed FC strength within the numerosity network could predict individual differences in the ANS in the left-out participant. We further examined whether the numerosity network could predict the performance of arithmetic skills in an independent dataset from the Human Connectome Project (HCP). Our results suggest that the summed FC strength within the numerosity network can specifically predict individual differences in arithmetic skills, rather than domain-general cognitive abilities.

Materials and methods

Overview

The ANS dataset consisted of a primary dataset and a validating dataset, which were used to obtain and verify the numerosity network. The HCP dataset was used to investigate whether the numerosity network identified in the ANS dataset could predict individual differences in arithmetic skills, language-comprehension ability, and other domain-general cognitive abilities. In the primary ANS dataset, we constructed a numerosity CPM based on behavior performance in an NC task and resting-state fMRI data, to predict the individual differences in ANS acuity, using a LOO cross-validation procedure. The reliability of the identified numerosity network was verified in an independent validating ANS dataset using the same NC task. Subsequently, in the HCP dataset, we examined whether the identified numerosity network could predict individual differences in arithmetic skills and other domain-general cognitive skills.

ANS dataset

Participants

The behavioral and fMRI data were collected from two groups of college students (Groups A and B). The data from Group A (154 participants; age, 20–25 (mean, 21.66; SD, 1.07) years; 65 males) were used to construct the numerosity CPM, whereas the data from Group B (145 participants; age, 20–25 (mean, 21.39; SD, 0.91) years; 36 males) were used to validate the reliability of the numerosity network. The Group A and Group B datasets corresponded to the ANS primary and validating datasets, respectively. None of the participants had a history of neurological disorder (e.g. mental retardation or traumatic brain injury) or psychiatric illness. The experimental protocol was approved by the Institutional Review Board of Beijing Normal University. Written informed consent was obtained from all participants before the study.

Nonsymbolic number-comparison task

Similar to previous research (Halberda et al. 2008), a classical NC paradigm was used to measure the ANS acuity of the participants, who completed 40 test trials after practicing five practice trials. In each trial, a spatially intermixed blue and yellow dot display was presented on a computer screen for 750 ms (Fig. 1A). The participants indicated which color was more numerous by pressing a key. Eleven different ratios between the two sets of colored dots were used, i.e. 12:11, 11:10, 10:9, 9:8, 8:7, 7:6, 6:5, 10:8, 8:6, 9:6, and 12:6. The color of the dots was counterbalanced across trials. The dot size of each ratio level was controlled: the average blue dot size variation was equal to the average yellow dot size variation.

The minimum distinguishable difference between the number of blue and yellow dots that produces a noticeable response was estimated as each participant's ANS acuity, which is known as the Weber fraction. The Weber

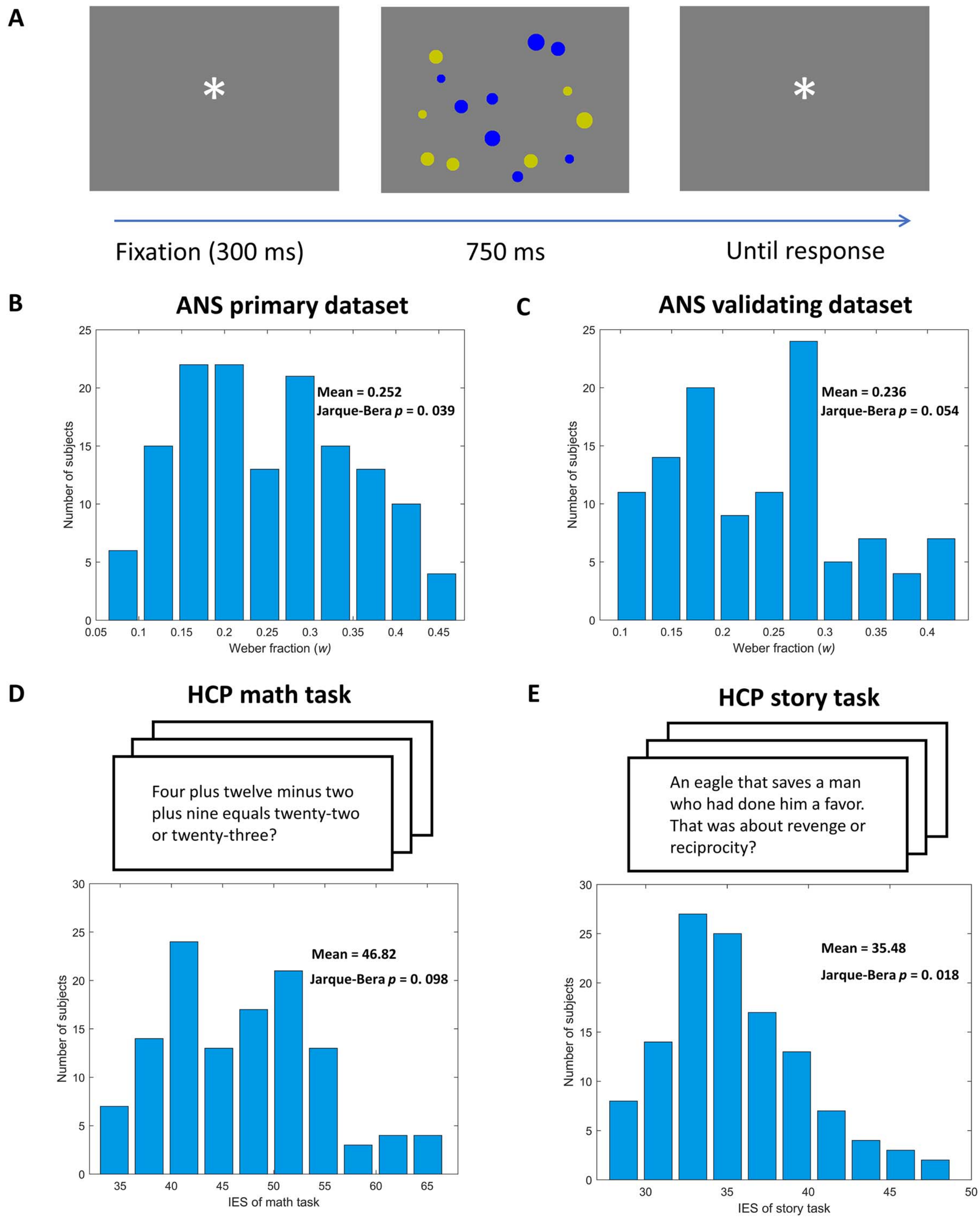

Fig. 1. Experimental task and group performance of the ANS and HCP story/math tasks. A) Paradigm of the NC task. B) Histogram of the Weber fraction (w), ANS acuity, for the ANS primary dataset ($n = 141$), as determined using a psychophysical model for each participant. C) Histogram of the Weber fraction (w) for the ANS validating dataset ($n = 112$). D and E) experimental task and histogram of IES performance for the HCP math/story tasks ($n = 117$). NC, number comparison; HCP, Human Connectome Project; ANS, approximate number system; IES, inverse efficiency score.

fraction of each participant was estimated using a QUEST routine (Watson and Pelli 1983; Baldassi and Burr 2000), which provides a given number of sequential trials and updates the probability distribution function (PDF) of the Weber fraction based on the participant's response and current PDF, according to Bayes' Rule. After the final trial, the mean PDF value was recorded as the participant's Weber fraction and was used to represent the participant's ANS acuity. Thus, a greater Weber fraction score corresponded to a poorer ANS acuity.

Participants with divergent estimating sequences were excluded by visual inspection. Moreover, participants with a Weber fraction greater than 0.5 were also excluded from subsequent analyses (e.g. a Weber fraction greater than 0.5 means an inability to distinguish 12 dots from 6 dots). In summary, 13 participants in the ANS primary dataset and 25 participants in the ANS validating dataset were excluded from further analysis.

MRI data and preprocessing

The structural and resting-state fMRI data were acquired using a Siemens 3 T Trio scanner (MAGENTOM Trio, a Tim system) with a 12-channel phased-array head coil. T1-weighted structure images were acquired with a magnetization-prepared rapid gradient-echo (MPRAGE) sequence (TR/TE/TI = 2.53 s/3.45 ms/1.1 s; FA = 7°; voxel size = 1 × 1 × 1 mm; slice thickness = 1.33 mm) for each participant. The resting-state data were acquired using a T2*-weighted GRE-EPI sequence (TR = 2000 ms; TE = 30 ms; flip angle = 90°; number of slices = 33; voxel size = 3.125 × 3.125 × 3.6 mm; number of volumes = 240).

Resting-state fMRI images were preprocessed using the FMRI Expert Analysis Tool (version 5.98), which is part of the FMRIB Software Library (www.fmrib.ox.ac.uk/fsl). The first four volumes of each participant were discarded to allow for the stabilization of magnetization. In addition to head motion correction, brain extraction, spatial smoothing (FWHM = 6 mm), grand-mean intensity normalization, and removal of a linear trend, several other preprocessing steps were used to reduce spurious variance. These steps included the use of a temporal band-pass filter (0.01–0.08 Hz) to retain low-frequency signals exclusively, and regression of the time course obtained from the motion-correction parameters, the mean signals of the cerebrospinal fluid and white matter, and the first derivatives of these signals. Subsequently, all functional images were aligned to the structural images using the linear image registration tools of FMRIB and warped to the MNI152 template using the nonlinear image registration tool of FMRIB. Because head motion might confound the FC analyses, in the validating dataset, we excluded eight participants who had a frame-to-frame head motion estimation greater than 0.15 (Finn et al. 2015). Finally, 141 participants were retained in the ANS primary dataset, and 112 participants were retained in the ANS validating dataset. There was no correlation between head motion and the Weber fraction in the retained participants in both datasets (primary dataset: $r = -0.015$, $P > 0.8$; validating dataset: $r = -0.007$, $P > 0.9$).

Functional connectivity calculation

FC was calculated between ROIs or "nodes" in Shen's atlas, which maximized the similarity of the time series of the voxels within each node (Shen et al. 2010, 2013). A subset of nodes of the 268-node functional brain atlas was used in our study because some resting fMRI scans did not cover the full brainstem and cerebellum. Thus, we focused on nodes in the neocortex by removing all nodes of the cerebellum, brainstem, and subcortex (67 nodes in total). The remaining 201 nodes are shown in Supplementary Fig. 1. The mean time course of each node was extracted as a measure of spontaneous neural activity in that node. Pearson's correlation coefficients (r) were calculated between the time courses of each pair of nodes and normalized using Fisher's r-to-Z transformation. Finally, a 201 × 201 FC matrix was obtained for each participant in the primary and validating datasets.

HCP dataset

Participants

The data used for evaluating the predictive power of the numerosity network regarding the individual differences in arithmetic skills, language comprehension, nonverbal WM, executive function/inhibition, and gF were from the Human Connectome Project (Van Essen et al. 2013) (Q1 and Q2 HCP data releases; 131 participants; age, 22–35 years; 30 males). Participants in HCP dataset were community samples with at least 9 years of education experience. Resting-state fMRI data from both left–right (LR) and right–left (RL) phase-encoding runs in day 1 (HCP filenames: rfMRI_REST1) were included in subsequent analyses. Consistent with the preprocessing procedure of the ANS dataset, we also excluded seven participants with a frame-to-frame head motion estimation greater than 0.15 (violated in either the LR or RL phase-encoding runs, HCP: Movement_RelativeRMS_mean). Measurements of arithmetic skills, language comprehension, nonverbal WM, executive function/inhibition, and gF were obtained for the remaining 124 participants simultaneously.

Behavioral test

Language (math vs. *story) task*

Arithmetic skills and language-comprehension abilities were assessed using a language (math/story) task in which the participants answered math-related and story-related questions after hearing auditory blocks. The HCP language task consists of two runs that interleave four blocks of a math task and four blocks of a story task (Binder et al. 2011). The math task engages the attention of participants continuously using mental arithmetic. The math task included auditory trials and required the participants to complete addition and subtraction problems, followed by a 2-alternative forced-choice task. For example, "Four plus twelve minus two plus nine equals *twenty-two* or *twenty-three*?" The math task is adaptive, to maintain a similar level of difficulty across participants. The story task included brief and engaging auditory stories (5–9 sentences). For example, after a story about an eagle that saves a man who had done him a favor, participants were asked the following question: "That was about *revenge* or *reciprocity*?" In both tasks, participants pushed a button to select either the first or the second choice. The median reaction time and accuracy were recorded in both the math and story tasks (HCP math task: Language_Task_Math_Acc and Language_Task_Math_Median_RT; HCP story task: Language_Task_Story_Acc and Language_Task_Story_Median_RT). We calculated the inverse efficiency score (IES) to integrate response time and accuracy

(Townsend and Ashby 1978). The IES was calculated as the median response time divided by the accuracy to evaluate arithmetic skills and language-comprehension abilities. There was no correlation between head motion and IES in the math/story task in the remaining set of 124 participants for math and story tasks (LR: IES of the math task, $r = 0.05$, $P > 0.5$; IES of the story task, $r = -0.03$, $P > 0.7$; RL: IES of the math task, $r = 0.01$, $P > 0.9$; IES of the story task, $r = -0.09$, $P > 0.3$).

Working-memory task

The nonverbal WM performance was measured using the visual-object N-back task with low (0-back) and high (2-back) conditions. The overall median reaction time (HCP: WM_Task_Median_RT) and accuracy (HCP: WM_Task_Acc) across all conditions in the WM task was recorded. To integrate response time and accuracy, we also calculated the IES, to represent the performance during the WM task. There was no correlation between head motion and IES in the WM task (LR: IES of the WM, $r = 0.10$, $P > 0.2$; RL: IES of the WM, $r = 0.05$, $P > 0.6$).

Flanker task

A flanker task was used to measure cognitive control (executive function/inhibitory control). This test required the participants to focus on a central stimulus while inhibiting attention to the stimuli flanking it. The central stimulus pointed either in the same direction as the flankers (congruent) or in the opposite direction (incongruent). Scoring was based on a combination of accuracy and reaction time. A larger flanker score indicates a better executive function/inhibitory control performance. In our research, both age-nonadjusted scale scores (HCP: Flanker_Unadj) and age-adjusted scale scores (HCP: Flanker_AgeAdj) were included. A larger Weber fraction value and IES in the math/story/WM correspond to a poorer ANS acuity or poor math-/story-related/WM performance. For easier comparisons with the other behavioral measures, we recoded the raw value of executive function/inhibition scores (Flanker_Unadj, Flanker_AgeAdj) by multiplying them by −1. Consequently, a larger value would also imply a worse task performance. There was no correlation between head motion and performance in the executive function/inhibition task (LR: Flanker_Unadj $r = 0.01$, $P > 0.8$; Flanker_AgeAdj $r = 0.01$, $P > 0.8$; RL: Flanker_Unadj $r = 0.04$, $P > 0.6$; Flanker_AgeAdj $r = 0.02$, $P > 0.8$).

Fluid intelligence

The fluid intelligence score was assessed using a form of Raven's progressive matrices with 24 items (the gF scores are integers indicating the number of correct items; HCP: PMAT24_A_CR). Similarly, we recoded the raw value of gF to error rate (ER), which was calculated as (24—gF)/24. Larger values would also imply a poorer behavior performance. There was no correlation between head motion and ER of gF (LR: ER of the gF, $r = 0.09$, $P > 0.3$; RL: ER of the gF, $r = 0.03$, $P > 0.7$).

MRI data and preprocessing

The structural and resting-state fMRI data in HCP dataset were acquired using a customized Siemens 3T scanner with a 32-channel Siemens receive head coil (Van Essen et al. 2013). T1-weighted structure images were acquired with a three-dimensional (3D) MPRAGE sequence (TR/TE/TI = 2.40 s/2.14 ms/1.0 s; FA = 7°; voxel size = 0.7 × 0.7 × 0.7 mm). The resting-state data were acquired using a T2*-weighted GRE-EPI sequence (TR = 720 ms; TE = 33.1 ms; flip angle = 52°; number of slices = 72; voxel size = 2.0 × 2.0 × 2.0 mm; number of volumes = 1,200).

The HCP minimal preprocessing pipeline was used for the HCP dataset (Glasser et al. 2013). This pipeline includes artifact removal, motion correction, and registration to the standard space. After this pipeline, several standard preprocessing procedures, including linear detrending, temporal filtering (0.01–0.1 Hz), regression of 12 motion parameters (HCP data; these included first derivatives, given as Movement_Regressors_dt.txt), and mean time courses of the white matter and CSF, as well as the global signal, were applied to the fMRI data (Finn et al. 2015). No spatial smoothing was used in HCP MRI dataset preprocessing.

Functional connectivity calculation

All steps used to construct resting-state FC networks were identical to those used for the ANS dataset. Data from both LR and RL phase-encoding runs were used to calculate connectivity matrices, respectively. The average of these two connectivity matrices was used as the connectivity matrix of the participants. Finally, a 201 × 201 FC matrix was obtained for each participant.

Numerosity CPM and numerosity network

Numerosity CPM

Numerosity CPM was established to predict individual differences in the ANS acuity using a LOO procedure (Rosenberg et al. 2016; Jangraw et al. 2018). One participant was left out for each LOO iteration, as the testing set. The remaining participants were regarded as the training set. Because the Jarque–Bera tests of normality revealed that these behavioral measures were not normally distributed (Fig. 1B–E, Supplementary Fig. 2A–D), we used Spearman's rank correlation for edge selection when constructing the numerosity CPM. Spearman's rank correlation was calculated across participants between the FC of each edge in the FC matrix and Weber fraction. Consistent with previous research (Rosenberg et al. 2016), for each LOO iteration, in the training set, the edges showing significant positive or negative correlation with a P value below a threshold of $P < 0.01$ across participants were included in a positive or negative network, respectively. The network strength of the training participant was defined as the participant's summed strength of all FCs within the positive or negative network. A general linear model (GLM) was fit to relate the summed FC strength to the Weber fraction for positive and negative networks, respectively. Subsequently, in the testing set, the GLM was used to obtain the predicted Weber fraction of the left-out participant from his/her summed FC strength within the positive and negative network separately. After the LOO procedure was repeated for each participant iteratively, we calculated the Pearson's correlation coefficient between the observed and predicted Weber fractions across all participants, to evaluate the predictive power. The standard deviation (SD) method was applied to the Weber fraction to remove participants with an outlier of the Weber fraction. The Weber fraction of each participant was within the range of ± three times the SD around the mean value. No participant was removed from the ANS primary dataset.

Permutation test

To confirm the reliability of the prediction results, we fixed the predicted Weber fraction and shuffled the observed Weber fraction across participants 10,000 times. We calculated the Pearson's correlation between the predicted and observed (randomly shuffled) scores. The number of times that the correlation coefficients in the set of 10,000 permutation tests outperformed the correlation coefficient from the numerosity CPM divided by 10,000 was recorded as the permutation test probability.

Numerosity network

Note that the positive or negative networks differed across iterations. To obtain a unique numerosity network, we generated a final numerosity network by selecting the overlapping edges across the network of all LOO iterations of the numerosity CPM (i.e. the overall network consisted of the edges found in all LOO iterations). Subsequently, the numerosity neuromarker was defined as the summed FC strength within the final numerosity network in this study. We evaluated the predictive power of the numerosity neuromarker for several datasets, including an independent ANS dataset and the HCP dataset. Note that we used the numerosity network to denote the final numerosity network.

Predictive power of the numerosity network

Predictive power of the numerosity network for the ANS validating dataset

The ANS validating dataset was used to confirm the reliable predictive power of the numerosity network for an independent dataset with the same NC task. To evaluate the predictive power of the numerosity network, we calculated the Pearson's correlation coefficients between the summed strength within the numerosity network and the Weber fraction of the ANS validating dataset. The standard deviation (SD) method (beyond the range of ± three times the SD around the mean value) was used to exclude participants with an outlier behavior performance from the correlation analysis. No participant was excluded from the ANS validating dataset.

To confirm the significance of the results obtained from the ANS validating dataset, we fixed the strength of the numerosity network and shuffled the observed Weber fraction 10,000 times and calculated the Pearson's correlation coefficients between them. The permutation test probability was calculated as the number of times that the correlation coefficients in a set of 10,000 permutation tests outperformed the predictive power of the numerosity network, divided by 10,000.

Predictive power of the numerosity network for the HCP dataset

The HCP dataset was used to evaluate the predictive ability of the numerosity network for arithmetic skills, language comprehension, and other domain-general cognitive skills (WM, executive function/inhibitory control, and gF). The predictive power of the numerosity network was evaluated by Pearson's correlation coefficients between the strength of the numerosity network and the behavioral performance on the math task, story tasks, and domain-general cognitive tasks. To explore whether the predictive ability of the numerosity network is specific to arithmetic skills, we also performed a Pearson's partial correlation analysis between summed FC strengths within the numerosity network and IESs in the math task, using age, gender, and domain-general cognitive skills (Flanker_Unadj, Flanker_AgeAdj, IES of the WM task, and ER of the gF test) as covariates of no interest. Two participants (participant IDs: 172332, 221319) in the HCP math task, two participants (participant IDs: 194140, 293748) in the HCP story task, and three participants (participant IDs: 169343, 214423, 579665) in the HCP WM task were excluded because their performances were beyond the range of ± three times the SD around the mean value. No participants in other control tasks were excluded because of outlier behavior performance. Hence, these seven participants were excluded from the correlation analysis. The permutation test procedures were the same as those used for the ANS validating dataset.

Anatomical distribution of the numerosity network

To determine the anatomical location of the numerosity network, we grouped the 201 nodes into 7 macroscale regions, including the prefrontal (PFC), motor (Mot), insular (Ins), parietal (Par), temporal (Tem), occipital (Occ), and limbic (Lim) cortices. We calculated the number of functional connections within or between all 7 macroscale regions. To investigate the hub nodes in the numerosity network, we ranked all nodes according to their degree centrality (DC; i.e. the number of direct functional connections between a given node and the remaining nodes within the numerosity network) (Wang, Jiao, et al. 2017b). The nodes and edges of the numerosity network were presented using BrainNet Viewer (V1.70, www.nitrc.org/projects/bnv) (Xia et al. 2013).

Table 1. Summary of the behavioral data.

Task	Metric	Participant #	Mean	SD	Range	Jarque–Bera (P)
ANS primary dataset	Weber fraction	141	0.252	0.099	[0.061, 0.474]	0.039
ANS validating dataset	Weber fraction	112	0.236	0.088	[0.094, 0.430]	0.054
HCP math	Inverse efficiency score	117	46.824	7.729	[32.72, 66.77]	0.098
HCP story	Inverse efficiency score	117	35.480	4.325	[27.58, 48.85]	0.018
HCP flanker	Unadjusted scale score	117	113.74	9.455	[92.73, 142.11]	0.049
	Age-adjusted scale score	117	104.54	9.097	[80.34, 123.56]	0.228
HCP working memory	Inverse efficiency score	117	10.248	1.911	[6.708, 15.827]	0.0074
HCP gF test	Error rate	117	0.294	0.185	[0, 0.7083]	0.018

Jarque–Bera tests indicated significant departures from normality, supporting the use of Spearman's rank correlations. Because we sought to identify any potential departures from normality, no correction for multiple comparisons was applied across these tests. All P values are based on two-tailed tests. ANS, approximate number system; HCP, Human Connectome Project.

Results

Behavioral performance

In the ANS dataset, similar to previous research (Halberda et al. 2008), a classical nonsymbolic NC task paradigm was used to measure the ANS acuity (Weber fraction) for each participant (Halberda et al. 2008). In the HCP dataset, we calculated the behavioral measures of the math task, story task, and other cognitive tasks for each participant, including the gF, nonverbal WM, and flanker tasks. The performance in these behavioral measures is shown in Table 1. In the ANS dataset, there was no significant difference in the ANS acuity between male and female participants ($t_{251} = 1.37$, $P > 0.17$).

The numerosity network can predict individual differences in ANS acuity

In the ANS primary dataset (141 participants), using the protocol of Shen et al. (Shen et al. 2017), we adopted the LOO cross-validation method to test whether the intrinsic FC profile can predict the Weber fraction. For the positive network, the correlation between the observed and predicted Weber fractions was significant ($r = 0.204$, $P < 0.015$), whereas for the negative network, this correlation was nonsignificant ($r = 0.032$, $P > 0.7$) (Fig. 2). We then conducted a permutation test to further confirm the reliability of these results. To this end, we shuffled the behavior performance across participants 10,000 times and calculated the Pearson's correlation coefficients between the observed (randomly shuffled) and predicted Weber fraction scores. We found that the positive network outperformed the set of 10,000 permutation tests ($P_{perm} < 0.009$). In comparison, the negative network did not outperform the set of 10,000 permutation tests ($P_{perm} > 0.3$). There was a significant difference in the predictive power between the positive and negative networks (Steiger's $z = 2.67$, $P = 7.6 \times 10^{-3}$) (Steiger 1980). This suggests that the strength of the FC profile within the positive networks could predict individual differences in ANS acuity. However, the negative network did not show such predictive ability. Therefore, in subsequent analyses, we focused on the predictive power of the positive network. Moreover, the predictive ability of the numerosity network could not be alternatively explained by head motion, as the average frame-to-frame motion was not correlated with the Weber fraction ($r = -0.015$, $P > 0.8$).

Note that the positive networks differed across iterations. We obtained a final numerosity network by selecting the overlapping edges of the positive networks across all iterations. Finally, the numerosity network included 87 nodes and 80 edges.

Subsequently, we evaluated whether the final numerosity network could successfully predict individual differences in ANS acuity in the independent validating dataset (112 participants). The correlation between the summed FC strength within the numerosity network and the Weber fraction across participants was significant ($r = 0.236$, $P < 0.013$) (Fig. 3A). A same procedure of permutation tests was applied. The numerosity network also outperformed the set of 10,000 permutations ($P_{perm} < 0.007$). This suggests that the final numerosity network identified here could serve as a connectome-based neuromarker to predict the ANS acuity of an individual reliably.

The summed FC strength within the numerosity network predicted arithmetic skills specifically

We further evaluated whether the summed FC strength within the numerosity network identified above could predict individual differences in arithmetic skills in an independent dataset, to address our second aim. The results of this analysis revealed that the summed FC strength within the numerosity network was significantly correlated with the inverse efficiency score (IES) (Townsend and Ashby 1978) in the math task ($r = 0.255$, $P < 0.005$) (Fig. 3B). The numerosity network also outperformed the set of 10,000 permutation tests in which we randomly shuffled the IES performance in the math task across participants ($P_{perm} < 0.002$). Even when age, gender, and domain-general cognitive skills (executive function, WM, and gF) were included as covariates of no interest, the Pearson's partial correlation between the summed FC strength within the numerosity network and the IESs in the math task remained significant ($r = 0.221$, $P < 0.03$, $P_{perm} < 0.004$). Moreover, a greater summed FC strength within the numerosity network was associated with a larger IES measure of the arithmetic skills. In turn, greater summed FC strength

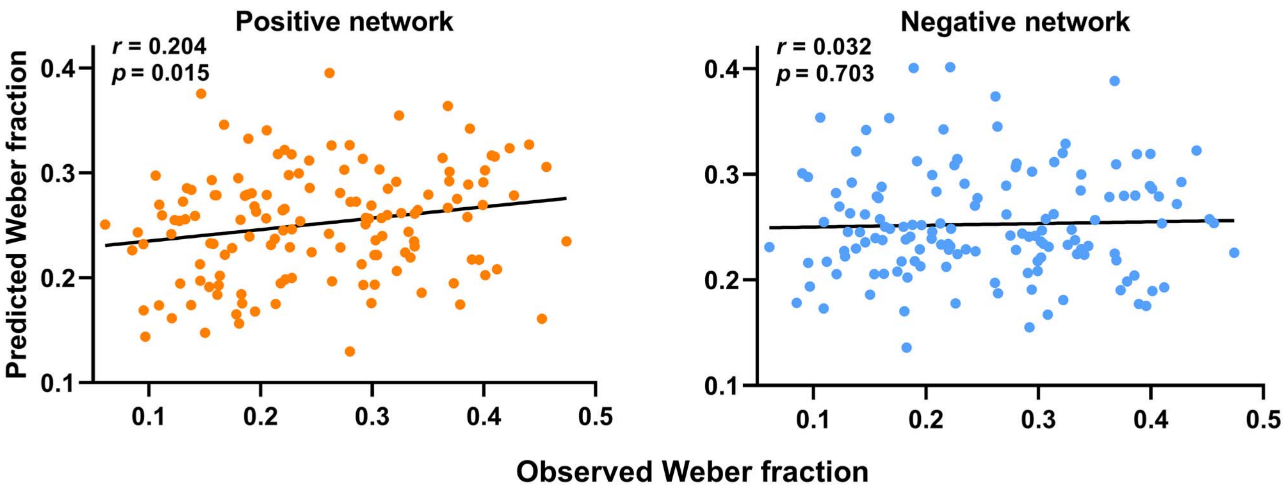

Fig. 2. CPMs predicted the Weber fraction. Scatter plots showing correlations between the observed Weber fraction and predictions obtained using the positive (left) and negative (right) networks. Network models were iteratively trained on behavior and MRI data from $n-1$ participants in the ANS primary dataset and tested on the left-out participant. The r and P values placed above each plot indicate the Pearson's correlation coefficient between the observed and predicted Weber fraction and the corresponding significance level, respectively. CPM, connectome-based predictive modeling; ANS, approximate number system.

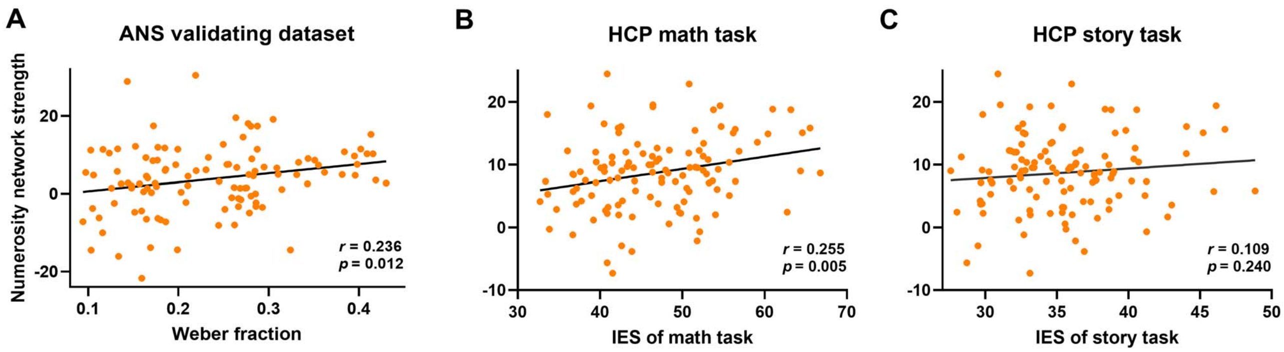

Fig. 3. Predictive power of the numerosity neuromarker. Pearson's correlations between the network strength of the numerosity network and behavior performance of Weber fraction in the ANS validating dataset (A), IES of math task (B), and IES of story task (C) in the HCP dataset were calculated to assess the predictive power of the numerosity neuromarker. The *r* and *P* values indicate the Pearson's correlation coefficient between the numerosity network strength and behavior performance and the corresponding significance level, respectively. ANS, approximate number system; IES, inverse efficiency score; HCP, Human Connectome Project.

within the numerosity network was related to a higher Weber fraction (i.e. poorer numerosity precision) in the NC task and a larger IES (i.e. poorer arithmetic skills) in the math task. This indicates that the numerosity network could also predict the performance of arithmetic skills, even in a completely new dataset.

To explore whether the predictive ability of the numerosity network is general for broader domain-general cognitive abilities, we next evaluated whether the numerosity network could predict performance on the story task and other domain-general cognitive tasks, which were shown in Table 2. We found that the summed FC strength within the numerosity network could not predict the IES performance in the story task (Fig. 3C), unadjusted and age-adjusted scale scores in the flanker task (Supplementary Fig. 3A and B), IES performance in the nonverbal WM task (Supplementary Fig. 3C), and the ER of the gF test (Supplementary Fig. 3D). The permutation tests confirmed that the numerosity network failed to outperform the set of 10,000 permutations (all $P_{\text{perm}} > 0.05$).

In summary, these findings suggest that the numerosity network could predict individual differences in arithmetic skills specifically, but not in language-comprehension abilities and other domain-general cognitive abilities, such as gF scores, nonverbal WM, and executive functioning.

Anatomical locations of the network edges

The numerosity network involved widespread regions in the neocortex (Fig. 4A and B). We then examined the anatomical locations of the network edges within the numerosity network. Contralateral connections (50 edges) were more common than ipsilateral connections (30 edges) in the numerosity network ($\chi^2(1) = 5.002$, $P < 0.03$). There were 19 right–right connections and 11 left–left connections in the numerosity network ($\chi^2(1) = 2.379$, $P > 0.1$) (Fig. 4C).

The numerosity network included many numerosity-related brain regions reported previously, such as the early visual cortex (DeWind et al. 2019), parietal sensory cortex (Lasne et al. 2019), angular gyrus (Klein et al. 2019),

Table 2. Predictive ability of the numerosity network on boarder domain-general cognitive abilities.

Domain-general tasks	Metric	Correlation with strength of numerosity network	
		r value	*P* value
HCP story	Inverse efficiency score	0.109	>0.2
HCP flanker	Unadjusted scale score	−0.047	>0.6
	Age-adjusted scale score	−0.037	>0.6
HCP working memory	Inverse efficiency score	0.149	>0.1
HCP gF test	Error rate	0.092	>0.3

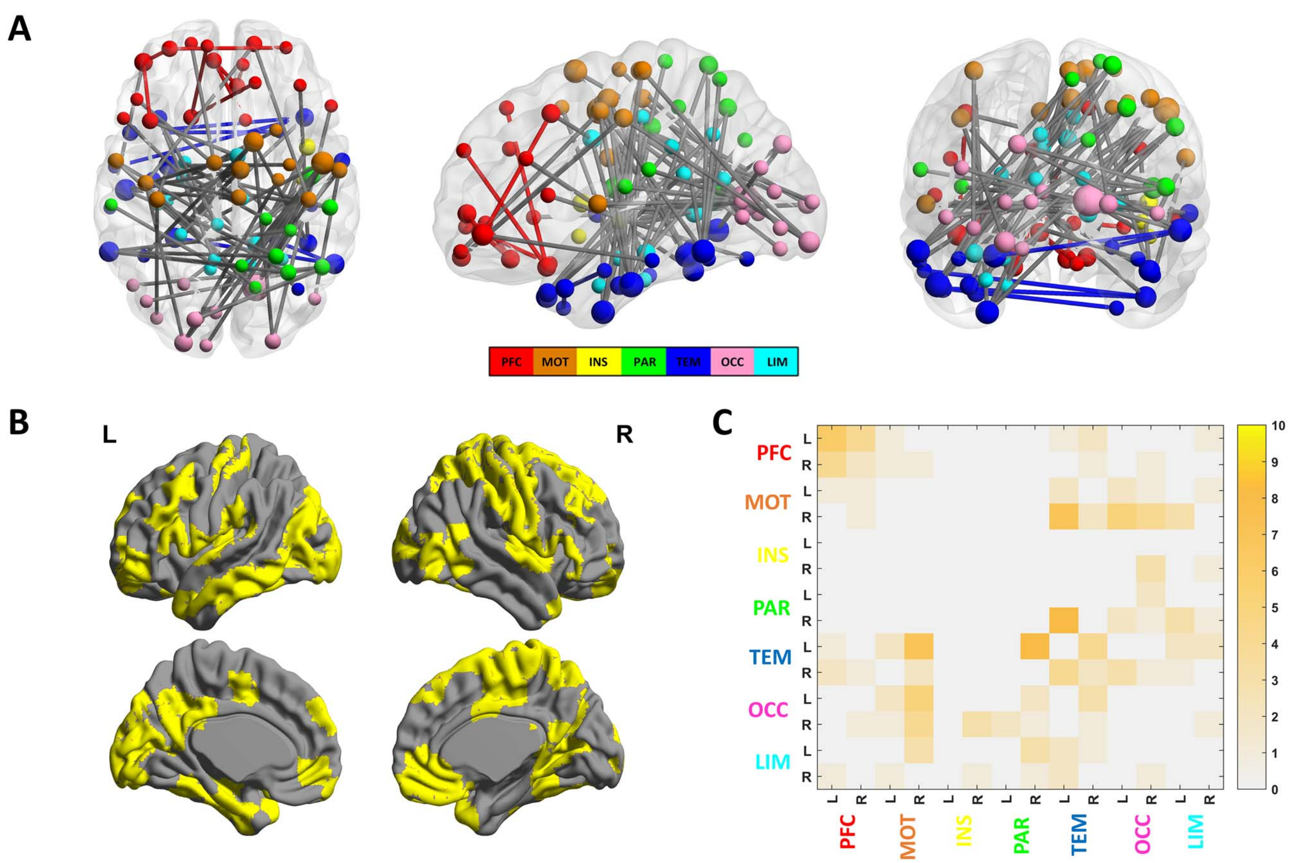

Fig. 4. Anatomical locations of the numerosity network. A) Three-dimensional view of the numerosity network edges and nodes on a glass brain. Nodes located in the cerebellum, subcortical, and brainstem areas were not included in our analysis. Hence, macroscale regions included the prefrontal cortex (PFC), motor cortex (Mot), insula (Ins), parietal (Par), temporal (Tem), occipital (Occ), and limbic (including the cingulate cortex, amygdala, and hippocampus; Lim) areas. A larger size indicates that the node had a greater DC value. B) Cortical areas of the numerosity network. C) All edges were grouped by macroscale region and hemisphere. The number of edges within or between the macroscale regions is shown. DC, degree centrality; L, left hemisphere, R, right hemisphere.

medial temporal gyrus (MTG) (Kutter et al. 2018), motor cortex (Anobile et al. 2021), and prefrontal cortex (Ramirez-Cardenas and Nieder 2019). The occipital (primary visual and visual association) nodes were mostly connected to parietal and motor regions (Fig. 5A). There were dense connections between left temporal nodes and the right motor and right parietal regions (Fig. 5B and C). The PFC nodes were mainly connected within the lobes (Fig. 5D). Detailed information of nodes (DC > 1) and remaining edges could be found in subsequent appendices (Supplementary Fig. 4, Supplementary Table 1).

Discussion

In the present study, using the CPM analysis, we obtained a numerosity network based on the numerosity comparison task, the summed FC strength of which reliably predicted individual differences in ANS acuity among individuals. This numerosity network consisted of functional connections between widely distributed brain regions, suggesting that a whole-brain, widely distributed numerosity network may serve as a neuromarker of an individual's nonsymbolic number acuity. Another interesting finding of our study was that the connectome-based neuromarker of ANS acuity identified using a numerosity comparison task also predicted individual differences in arithmetic skills specifically, suggesting that numerosity perception and arithmetic skills share a similar functional neural basis. However, the summed FC strength within the numerosity network did not predict individual differences in other domain-general abilities, i.e. nonverbal WM, executive control, fluid intelligence,

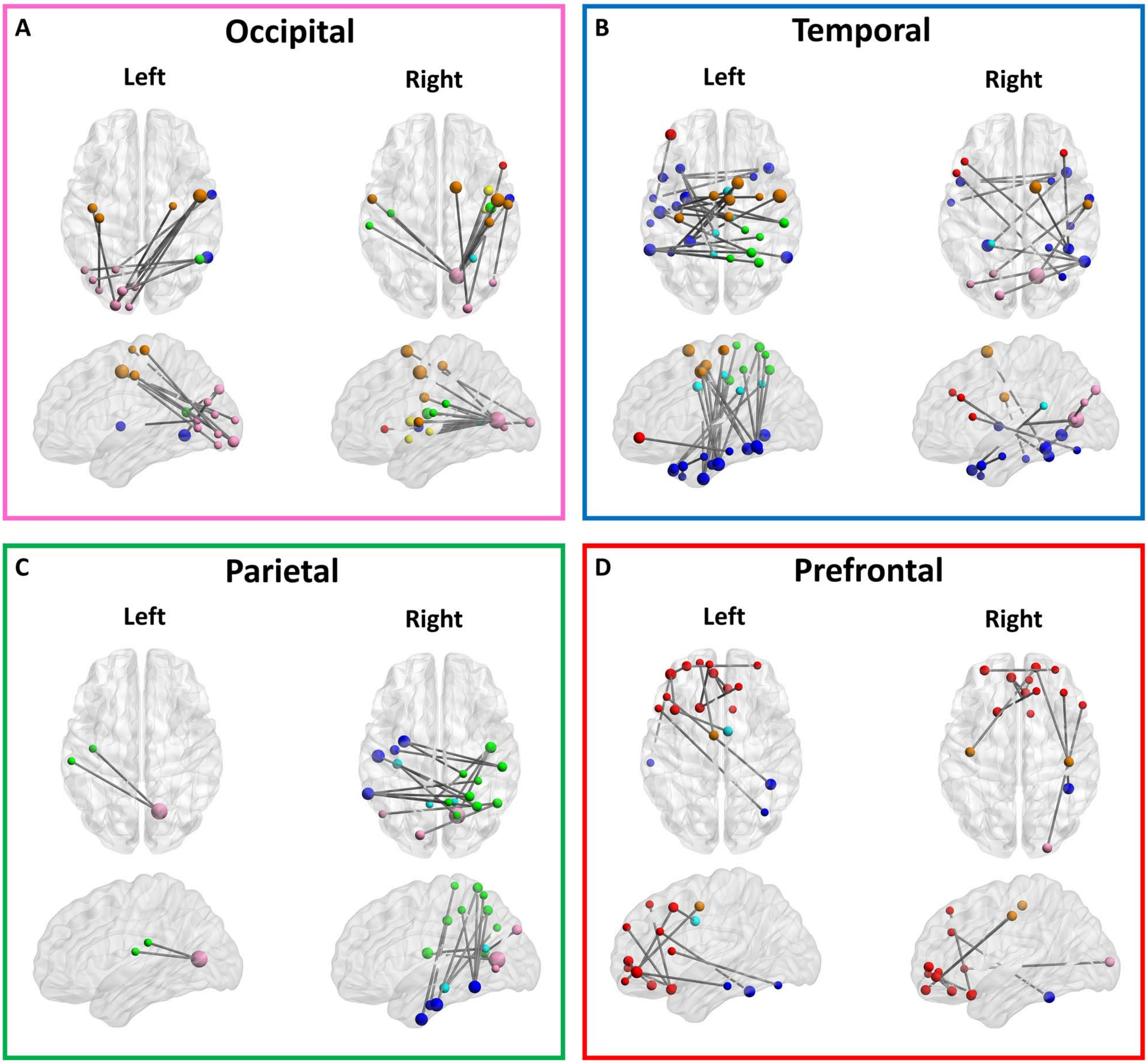

Fig. 5. Functional connections in the numerosity network in each of four macroscale regions: A) occipital, B) temporal, C) parietal, and D) prefrontal regions. The nodes in left or right hemisphere of each macroscale region were shown respectively, indicated with "left" and "right" labels. The indication of each macroscale region is consistent with Fig. 4.

and language comprehension, indicating the specificity of the association between the numerosity network and arithmetic skills.

Our study used the CPM analysis to examine the relationship between individual differences in functional brain connectivity and individual differences in ANS acuity. Unlike the traditional methods used for establishing the brain–behavior relationships (e.g. correlation or simple regression), which might overfit the data and sometimes fail to generalize to novel data, the CPM analysis adopts a cross-validation approach to eliminate these spurious effects effectively (Shen et al. 2017), which renders it a reliable solution for the development of a neuroimaging-based biomarker of ANS acuity. It is possible that other multivariate predictive models, e.g. support vector regression (SVR) method, may outperform CPM in terms of prediction accuracy. However, the major advantage of the CPM approach is that the obtained predictive networks can be clearly interpreted.

The numerosity CPM identified here involved multiple widespread regions reported in previous researches of numerosity processing, including visual cortex (Park et al. 2016; Fornaciai et al. 2017), sensorimotor cortex (Sawamura et al. 2002; Piazza et al. 2006; Anobile et al. 2021), parietal cortex(Piazza et al. 2007; Arsalidou and Taylor 2011), MTL (Kutter et al. 2018), and lateral prefrontal cortex (Nieder et al. 2002). Therefore, the numerosity representation might rely on the cooperative work of multiple widespread brain regions. Moreover, the numerosity network included many connections between occipital and sensorimotor nodes, indicating the interaction between visual perception, action planning, and execution in the numerosity system (Anobile et al. 2021). In addition, the numerosity network showed dense connections between the left temporal and right premotor, and between the left temporal and right parietal sensorimotor regions, which may represent a functional coupling between the ventral and dorsal areas in the numerosity system (Yang et al. 2020). Taken together, our findings provided a complementary measure of numerosity processing and arithmetic skills to the traditional approaches that focus on the specific function of discrete brain regions.

In our study, we found that an increased strength of the numerosity network was indicative of poor numerosity acuity or arithmetic skills. One possible explanation for this finding was that it might reflect a compensatory mechanism (Voets et al. 2009). Participants with a poor numerosity acuity may need more sensory representation of the stimulus or additional cognitive resources, which is consistent with the observation of hyperconnectivity in children with mathematical difficulties (Cappelletti and Price 2014; Rosenberg-Lee et al. 2014; Leibovich et al. 2017; Starr et al. 2017; Castaldi et al. 2018; Piazza et al. 2018; Wilkey and Price 2019). However, the current findings did not provide direct evidence of this phenomenon. Thus, future investigation is needed to test this hypothesis directly.

It should be noted that the association between nonverbal number acuity and symbolic math performance is the subject of a heated debate in this field (Xenidou-Dervou, De Smedt, et al. 2013a; Leibovich-Raveh et al. 2017; Schneider et al. 2017). Several studies have found a positive correlation between numerosity precision and math ability (Halberda et al. 2008; Starr et al. 2013) and proposed that the number sense may serve as a "start-up" tool for mathematics acquisition (Piazza 2010). However, other researchers argued that the correlation between nonverbal number acuity and math ability was usually weak (Desoete et al. 2010; Sasanguie et al. 2013; Chen and Li 2014; Fazio et al. 2014; Schneider et al. 2017), or even absent (Inglis et al. 2011; Fuhs and McNeil 2013; Gilmore et al. 2013). Therefore, they alternatively proposed that various domain-general abilities (e.g. executive functioning, WM, and knowledge of Arabic numerals) play critical roles in arithmetic development (Kolkman et al. 2013; Xenidou-Dervou, De Smedt, et al. 2013a; Xenidou-Dervou, Van Lieshout, et al. 2013b; Bull and Lee 2014; Göbel et al. 2014; Szűcs et al. 2014; Van Dooren and Inglis 2015). In the present study, we found that the numerosity network identified using the nonsymbolic NC task in one group of participants predicted the arithmetic skills in a novel group of participants, even when controlling for several domain-general cognitive abilities (i.e. language comprehension, nonverbal WM, executive control, and fluid intelligence). However, our results did not provide direct evidence in support of any of the two hypotheses because of methodological limitations. First, although we found that the numerosity network did not predict other domain-general abilities, we cannot rule out the possibility that the domain-general abilities themselves can predict arithmetic skills. Future investigation using the CPMs specifically constructed for domain-general abilities will help address this possibility. Second, most of the previous studies focused on the relationship between ANS acuity and arithmetic skills in behavior. In contrast, our study focused on the brain–behavior relationships and generated predictions of behavioral measures in novel participants based on their FCs. If there is an association between nonsymbolic acuity and math performance, our findings may provide a neural basis for this link. Nevertheless, our results suggest that numerosity perception and arithmetic skills share a similar functional basis and that ANS acuity provides a unique contribution to arithmetic skills.

The present study had several limitations. First, it is possible that other multivariate predictive models, e.g. SVR method, may slightly outperform CPM in terms of prediction accuracy (Shen et al. 2017). The major advantage of the CPM approach is that the obtained predictive networks can be clearly interpreted. Considering the trade-off between interpretability and prediction accuracies, we used the CPM in the current study. It would be helpful if future investigations could use other advanced predictive models and systematically compare their prediction accuracies and with that of the CPM approach. Second, the current investigation was carried out in adults. Thus, whether the same relationship holds in children remains an open question; moreover, if so, how does it change during development? Previous research has revealed that the precision of numerical acuity sharpens with age and the acquisition of formal mathematical education, which may reflect an ability to focus on numerical information while filtering out nonnumerical information (Piazza et al. 2013; Starr et al. 2017; Castaldi et al. 2018; Wilkey and Price 2019). Future investigation testing the change in the numerosity network associated with the within-individual changes in numerosity perception and mathematical abilities over the years can inform on the neural mechanism underlying these processes and their relationships during development. Third, we did not include an investigation of symbolic numbers, which is most closely related to the learning of mathematics. Further research could use symbolic and nonsymbolic stimuli to address the common and distinct functional architectures of these tasks related to number perception and mathematical abilities. Fourth, because there were no available measures of the number sense in the HCP dataset and arithmetic performance in our dataset, it is possible that the number sense and arithmetic skills of participants in our dataset might be better than those in HCP participants. Although the education level of the participants in two datasets was compatible, future research with both measures in a single dataset is needed to clarify this issue.

In summary, our findings showed that the identified numerosity network might serve as an applicable, neuroimaging-based biomarker of nonverbal number-acuity and arithmetic skills.

Acknowledgments

We are indebted to Wenqiang Xu and Rongkang Deng for their valuable suggestions.

Supplementary material

Supplementary material is available at *Cerebral Cortex Journal* online.

Funding

This work was supported by National Science and Technology Innovation 2030 Major Program (2021ZD0203803), National Key R&D Program of China (2019YFA0709503), National Nature Science Foundation of China grant (61971289), Shenzhen Science and Technology Research Funding Program (JCYJ20170412111316339), Fundamental Research Funds for the Central Universities, and Shenzhen-Hong Kong Institute of Brain Science-Shenzhen Fundamental Research Institutions.

Conflict of interest statement. The authors have no conflicts of interest.

Data availability

Each participant's functional matrix and behavior performance from the ANS and HCP math/story datasets are available online from https://github.com/Dzhang1989z/Numerosity-CPM. Raw data from the ANS dataset are available from the corresponding author upon reasonable request. Raw fMRI data from HCP are publicly available by application (https://www.humanconnectome.org).

Code availability

Custom codes are available online from https://github.com/Dzhang1989z/Numerosity-CPM.

References

Agrillo C, Parrish AE, Beran MJ. How illusory is the solitaire illusion? Assessing the degree of misperception of numerosity in adult humans. *Front Psychol*. 2016:7:1663.

Anobile G, Arrighi R, Castaldi E, Burr DC. A sensorimotor Numerosity system. *Trends Cogn Sci*. 2021:25:24–36.

Arsalidou M, Taylor MJ. Is 2 + 2 = 4? Meta-analyses of brain areas needed for numbers and calculations. *NeuroImage*. 2011:54: 2382–2393.

Baldassi S, Burr DC. Feature-based integration of orientation signals in visual search. *Vis Res*. 2000:40:1293–1300.

Barth H, Kanwisher N, Spelke E. The construction of large number representations in adults. *Cognition*. 2003:86:201–221.

Beaty RE, Kenett YN, Christensen AP, Rosenberg MD, Benedek M, Chen Q, Fink A, Qiu J, Kwapil TR, Kane MJ, et al. Robust prediction of individual creative ability from brain functional connectivity. *Proc Natl Acad Sci U S A*. 2018:115:1087–1092.

Binder JR, Gross WL, Allendorfer JB, Bonilha L, Chapin J, Edwards JC, Grabowski TJ, Langfitt JT, Loring DW, Lowe MJ, et al. Mapping anterior temporal lobe language areas with fMRI: a multicenter normative study. *NeuroImage*. 2011:54:1465–1475.

Biswal B, Yetkin FZ, Haughton VM, Hyde JS. Functional connectivity in the motor cortex of resting human brain using echo-planar MRI. *Magn Reson Med*. 1995:34:537–541.

Brankaer C, Ghesquière P, De Smedt B. The development of numerical magnitude processing and its association with working memory in children with mild intellectual disabilities. *Res Dev Disabil*. 2013:34:3361–3371.

Bull R, Lee K. Executive functioning and mathematics achievement. *Child Dev Perspect*. 2014:8:36–41.

Cappelletti M, Price CJ. Residual number processing in dyscalculia. *NeuroImage Clin*. 2014:4:18–28.

Castaldi E, Mirassou A, Dehaene S, Piazza M, Eger E. Asymmetrical interference between number and item size perception provides evidence for a domain specific impairment in dyscalculia. *PLoS One*. 2018:13:e0209256.

Chen Q, Li J. Association between individual differences in non-symbolic number acuity and math performance: a meta-analysis. *Acta Psychol*. 2014:148:163–172.

Dehaene S, Cohen L. Cerebral pathways for calculation: double dissociation between rote verbal and quantitative knowledge of arithmetic. *Cortex*. 1997:33:219–250.

Dehaene S, Piazza M, Pinel P, Cohen L. Three parietal circuits for number processing. *Cogn Neuropsychol*. 2003:20:487–506.

Desoete A, Ceulemans A, Weerdt F, Pieters S. Can we predict mathematical learning disabilities from symbolic and non-symbolic comparison tasks in kindergarten? Findings from a longitudinal study. *Br J Educ Psychol*. 2010:82:64–81.

DeWind NK, Park J, Woldorff MG, Brannon EM. Numerical encoding in early visual cortex. *Cortex*. 2019:114:76–89.

Fazio LK, Bailey DH, Thompson CA, Siegler RS. Relations of different types of numerical magnitude representations to each other and to mathematics achievement. *J Exp Child Psychol*. 2014:123: 53–72.

Feigenson L, Dehaene S, Spelke E. Core systems of number. *Trends Cogn Sci*. 2004:8:307–314.

Finn ES, Shen X, Scheinost D, Rosenberg MD, Huang J, Chun MM, Papademetris X, Constable RT. Functional connectome fingerprinting: identifying individuals using patterns of brain connectivity. *Nat Neurosci*. 2015:18:1664–1671.

Fornaciai M, Brannon EM, Woldorff MG, Park J. Numerosity processing in early visual cortex. *NeuroImage*. 2017:157:429–438.

Fuhs MW, McNeil NM. ANS acuity and mathematics ability in preschoolers from low-income homes: contributions of inhibitory control. *Dev Sci*. 2013:16:136–148.

Geary DC, Moore AM. In: Cappelletti M, Fias W, editors. *Chapter 4—cognitive and brain systems underlying early mathematical development*. Amsterdam, Netherlands: Progress in Brain Research Elsevier; 2016. pp. 75–103

Gilmore C, Attridge N, Clayton S, Cragg L, Johnson S, Marlow N, Simms V, Inglis M. Individual differences in inhibitory control, not non-verbal number acuity, correlate with mathematics achievement. *PLoS One*. 2013:8:e67374.

Glasser MF, Sotiropoulos SN, Wilson JA, Coalson TS, Fischl B, Andersson JL, Xu J, Jbabdi S, Webster M, Polimeni JR, et al. The minimal preprocessing pipelines for the human connectome project. *NeuroImage*. 2013:80:105–124.

Göbel S, Watson S, Lervåg A, Hulme C. *Children's arithmetic development it is number knowledge, not the approximate number sense, that counts*. Newbury Park, California, USA: Psychological Science; 2014. p. 25.

Goense J, Whittingstall K, Logothetis N. Neural and BOLD responses across the brain. *Wiley Interdiscip Rev Cogn Sci*. 2012:3:75–86.

Gross HJ, Pahl M, Si A, Zhu H, Tautz J, Zhang S. Number-based visual generalisation in the honeybee. *PLoS One*. 2009:4:e4263.

Halberda J, Mazzocco MM, Feigenson L. Individual differences in non-verbal number acuity correlate with maths achievement. *Nature*. 2008:455:665–668.

Harvey BM. Quantity cognition: numbers, numerosity, zero and mathematics. *Curr Biol*. 2016:26:R419–R421.

Harvey BM, Ferri S, Orban GA. Comparing parietal quantity-processing mechanisms between humans and macaques. *Trends Cogn Sci*. 2017:21:779–793.

Hawes Z, Sokolowski HM, Ononye CB, Ansari D. Neural underpinnings of numerical and spatial cognition: an fMRI meta-analysis of brain regions associated with symbolic number, arithmetic, and mental rotation. *Neurosci Biobehav Rev*. 2019:103: 316–336.

He L, Zhou K, Zhou T, He S, Chen L. Topology-defined units in numerosity perception. *Proc Natl Acad Sci U S A*. 2015:112: E5647–E5655.

Inglis M, Attridge N, Batchelor S, Gilmore C. Non-verbal number acuity correlates with symbolic mathematics achievement: but only in children. *Psychon Bull Rev*. 2011:18:1222–1229.

Jangraw DC, Gonzalez-Castillo J, Handwerker DA, Ghane M, Rosenberg MD, Panwar P, Bandettini PA. A functional connectivity-based neuromarker of sustained attention generalizes to predict recall in a reading task. *NeuroImage*. 2018:166:99–109.

Kim G, Jang J, Baek S, Song M, Paik S-B. Visual number sense in untrained deep neural networks. *Sci Adv*. 2021:7:eabd6127.

Klein E, Willmes K, Bieck SM, Bloechle J, Moeller K. White matter neuro-plasticity in mental arithmetic: changes in hippocampal connectivity following arithmetic drill training. *Cortex*. 2019:114:115–123.

Kolkman ME, Kroesbergen EH, Leseman PPM. Early numerical development and the role of non-symbolic and symbolic skills. *Learn Instr*. 2013:25:95–103.

Kutter EF, Bostroem J, Elger CE, Mormann F, Nieder A. Single neurons in the human brain encode numbers. *Neuron*. 2018:100:753–761.e754.

Lasne G, Piazza M, Dehaene S, Kleinschmidt A, Eger E. Discriminability of numerosity-evoked fMRI activity patterns in human intra-parietal cortex reflects behavioral numerical acuity. *Cortex*. 2019:114:90–101.

Leibovich T, Katzin N, Harel M, Henik A. From “sense of number” to “sense of magnitude”: the role of continuous magnitudes in numerical cognition. *Behav Brain Sci*. 2017:40:e164.

Leibovich-Raveh T, Katzin N, Salti M, Henik A. Toward an integrative approach to numerical cognition. *Behav Brain Sci*. 2017:40:e194.

Libertus ME, Feigenson L, Halberda J. Is approximate number precision a stable predictor of math ability? *Learn Individ Differ*. 2013:25:126–133.

Lukowski SL, Rosenberg-Lee M, Thompson LA, Hart SA, Willcutt EG, Olson RK, Petrill SA, Pennington BF. Approximate number sense shares etiological overlap with mathematics and general cognitive ability. *Dermatol Int*. 2017:65:67–74.

Nasr K, Viswanathan P, Nieder A. Number detectors spontaneously emerge in a deep neural network designed for visual object recognition. *Sci Adv*. 2019:5:eaav7903.

Nieder A. Supramodal numerosity selectivity of neurons in primate prefrontal and posterior parietal cortices. *Proc Natl Acad Sci*. 2012:109:11860.

Nieder A. The neuronal code for number. *Nat Rev Neurosci*. 2016: 17:366–382.

Nieder A. Evolution of cognitive and neural solutions enabling numerosity judgements: lessons from primates and corvids. *Philos Trans R Soc London B Biol Sci*. 2017:373:20160514.

Nieder A, Miller EK. A parieto-frontal network for visual numerical information in the monkey. *Proc Natl Acad Sci*. 2004:101:7457–7462.

Nieder A, Freedman DJ, Miller EK. Representation of the quantity of visual items in the primate prefrontal cortex. *Science (New York, NY)*. 2002:297:1708–1711.

Park J, Dewind N, Woldorff M, Brannon E. Rapid and direct encoding of numerosity in the visual stream. *Cerebral cortex (New York, NY: 1991)*. 2015:26:e185.

Park J, DeWind NK, Woldorff MG, Brannon EM. Rapid and direct encoding of Numerosity in the visual stream. *Cereb Cortex*. 2016:26:748–763.

Piazza M. Neurocognitive start-up tools for symbolic number representations. *Trends Cogn Sci*. 2010:14:542–551.

Piazza M, Izard V, Pinel P, Le Bihan D, Dehaene S. Tuning curves for approximate numerosity in the human intraparietal sulcus. *Neuron*. 2004:44:547–555.

Piazza M, Mechelli A, Price CJ, Butterworth B. Exact and approximate judgements of visual and auditory numerosity: an fMRI study. *Brain Res*. 2006:1106:177–188.

Piazza M, Pinel P, Le Bihan D, Dehaene S. A magnitude code common to numerosities and number symbols in human intraparietal cortex. *Neuron*. 2007:53:293–305.

Piazza M, Pica P, Izard V, Spelke ES, Dehaene S. Education enhances the acuity of the nonverbal approximate number system. *Psychol Sci*. 2013:24:1037–1043.

Piazza M, De Feo V, Panzeri S, Dehaene S. Learning to focus on number. *Cognition*. 2018:181:35–45.

Pica P, Lemer C, Izard V, Dehaene S. Exact and approximate arithmetic in an Amazonian indigene group. *Science*. 2004:306:499–503.

Piffer L, Agrillo C, Hyde DC. Small and large number discrimination in guppies. *Anim Cogn*. 2012:15:215–221.

Ramirez-Cardenas A, Nieder A. Working memory representation of empty sets in the primate parietal and prefrontal cortices. *Cortex*. 2019:114:102–114.

Rosenberg MD, Finn ES, Scheinost D, Papademetris X, Shen X, Constable RT, Chun MM. A neuromarker of sustained attention from whole-brain functional connectivity. *Nat Neurosci*. 2016:19: 165–171.

Rosenberg-Lee M, Ashkenazi S, Chen T, Young C, Geary D, Menon V. Brain hyper-connectivity and operation-specific deficits during arithmetic problem solving in children with developmental dyscalculia. *Dev Sci*. 2014:18:351–372.

Sasanguie D, Göbel SM, Moll K, Smets K, Reynvoet B. Approximate number sense, symbolic number processing, or number–space mappings: what underlies mathematics achievement? *J Exp Child Psychol*. 2013:114:418–431.

Sawamura H, Shima K, Tanji J. Numerical representation for action in the parietal cortex of the monkey. *Nature*. 2002:415:918–922.

Schneider M, Beeres K, Coban L, Merz S, Schmidt S, Stricker J, De Smedt B. Associations of non-symbolic and symbolic numerical magnitude processing with mathematical competence: a meta-analysis. *Dev Sci*. 2017:20:e12372.

Shen X, Papademetris X, Constable RT. Graph-theory based parcellation of functional subunits in the brain from resting-state fMRI data. *NeuroImage*. 2010:50:1027–1035.

Shen X, Tokoglu F, Papademetris X, Constable RT. Groupwise whole-brain parcellation from resting-state fMRI data for network node identification. *NeuroImage*. 2013:82:403–415.

Shen X, Finn ES, Scheinost D, Rosenberg MD, Chun MM, Papademetris X, Constable RT. Using connectome-based predictive modeling to predict individual behavior from brain connectivity. *Nat Protoc*. 2017:12:506–518.

Smith SM, Vidaurre D, Beckmann CF, Glasser MF, Jenkinson M, Miller KL, Nichols TE, Robinson EC, Salimi-Khorshidi G, Woolrich MW, et al. Functional connectomics from resting-state fMRI. *Trends Cogn Sci*. 2013:17:666–682.

Starr A, Libertus ME, Brannon EM. Number sense in infancy predicts mathematical abilities in childhood. *Proc Natl Acad Sci*. 2013:110:18116.

Starr A, DeWind NK, Brannon EM. The contributions of numerical acuity and non-numerical stimulus features to the development of the number sense and symbolic math achievement. *Cognition*. 2017:168:222–233.

Steiger JH. Tests for comparing elements of a correlation matrix. *Psychol Bull*. 1980:87:245.

Szűcs D, Devine A, Soltesz F, Nobes A, Gabriel F. Cognitive components of a mathematical processing network in 9-year-old children. *Dev Sci*. 2014:17:506–524.

Townsend J, Ashby FG. Methods of modeling capacity in simple processing systems. *Cogn Theory*. 1978:3:200–239.

Van Dooren W, Inglis M. Inhibitory control in mathematical thinking, learning and problem solving: a survey. *ZDM*. 2015:47:713–721.

Van Essen D, Smith S, Barch D, Behrens T, Yacoub E, Ugurbil K. The WU-Minn human connectome project: an overview. *NeuroImage*. 2013:80:62–79.

Viswanathan P, Nieder A. Neuronal correlates of a visual "sense of number" in primate parietal and prefrontal cortices. *Proc Natl Acad Sci*. 2013:110:11187.

Voets NL, Adcock JE, Stacey R, Hart Y, Carpenter K, Matthews PM, Beckmann CF. Functional and structural changes in the memory network associated with left temporal lobe epilepsy. *Hum Brain Mapp*. 2009:30:4070–4081.

Wang J, Halberda J, Feigenson L. Approximate number sense correlates with math performance in gifted adolescents. *Acta Psychol*. 2017a:176:78–84.

Wang X, Jiao D, Zhang X, Lin X. Altered degree centrality in childhood absence epilepsy: a resting-state fMRI study. *J Neurol Sci*. 2017b:373:274–279.

Watson AB, Pelli DG. Quest: a Bayesian adaptive psychometric method. *Percept Psychophys*. 1983:33:113–120.

Wilkey ED, Price GR. Attention to number: the convergence of numerical magnitude processing, attention, and mathematics in the inferior frontal gyrus. *Hum Brain Mapp*. 2019:40:928–943.

Xenidou-Dervou I, De Smedt B, van der Schoot M, van Lieshout ECDM. Individual differences in kindergarten math achievement: the integrative roles of approximation skills and working memory. *Learn Individ Differ*. 2013a:28:119–129.

Xenidou-Dervou I, Van Lieshout E, Schoot M. Working memory in nonsymbolic approximate arithmetic processing: a dual-task study with preschoolers. *Cogn Sci*. 2013b:38:101–127.

Xia M, Wang J, He Y. BrainNet viewer: a network visualization tool for human brain connectomics. *PLoS One*. 2013:8:e68910.

Xu F, Spelke ES. Large number discrimination in 6-month-old infants. *Cognition*. 2000:74:B1–B11.

Yang H, He C, Han Z, Bi Y. Domain-specific functional coupling between dorsal and ventral systems during action perception. *Sci Rep*. 2020:10:21200.

Zhou C, Xu W, Liu Y, Xue Z, Chen R, Zhou K, Liu J. Numerosity representation in a deep convolutional neural network. *J Pac Rim Psychol*. 2021:15:18344909211012613.