國家科學及技術委員會補助專題研究計畫報告

以磁振造影研究探討數學能力之性別差異

報 告 類 別 ： 成果報告
計 畫 類 別 ： 個別型計畫
計 畫 編 號 ： MOST 110-2629-H-004-001-
執 行 期 間 ： 110年08月01日至111年10月31日
執 行 單 位 ： 國立政治大學心理學系

計畫主持人： 張葶葶

計畫參與人員： 學士級-專任助理：范惠君
碩士班研究生-兼任助理：陳心喻
碩士班研究生-兼任助理：黃如君
大專生-兼任助理：朱莫亞
大專生-兼任助理：程妍慈
大專生-兼任助理：劉奕姍
博士班研究生-兼任助理：伍贊達
博士後研究-博士後研究：陳乃鳳

報 告 附 件 ： 出席國際學術會議心得報告

本研究具有政策應用參考價值：■否 □是，建議提供機關
（勾選「是」者，請列舉建議可提供施政參考之業務主管機關）
本研究具影響公共利益之重大發現：□否 □是

中 華 民 國 112 年 01 月 12 日

中 文 摘 要 ： 一般大眾常認為數學能力有性別差異，且男性表現優於女性，然而實際文獻所報告之結果並不一致，且文獻間存在相當大之差異。根據TIMSS 1999、2003、2007及PISA 2006之成績統計，本國之四年級與八年級學生在科學與數學成績均無顯著之性別差異。過往雖不乏以大數據探討性別差異之行為與神經機制議題之研究，然直接檢驗數學認知下性別差異的磁振造影研究迄今為止僅有四篇文獻，且結果並不一致。為檢驗此議題，本研究以行為實驗及認知能力評估檢驗數學認知與學習之性別差異，並結合磁振造影實驗以檢驗數學認知與學習性別差異之神經機制，探討大腦結構與功能在進行算術作業時是否有性別差異，結果發現即便在行為表現上並無差異，女性卻在計算較複雜問題時前額葉及後頂葉活化比起男性強，使用機器學習模型得到此前額葉後頂葉之活化情形能夠預測，顯示兩性各自用各自的方式在進行相同的算術功能。本研究提供以教育、心理學、及認知神經科學領域探討男性及女性認知功能生理機制之實徵證據。

中 文 關 鍵 詞 ： 生理性別、心算、數學認知、磁振造影、頂葉、前額葉

英 文 摘 要 ： Numerous empirical studies have reported that males and females perform equally well in mathematical achievement. However, still to date, very limited is understood about the neural mechanisms of whether and how men and women demonstrate differences when solving mathematical problems. The present study aimed to tackle this issue by manipulating arithmetic problem complexity and investigating functional significance using fMRI in young adults. Participants were instructed to complete two runs of simple calculation tasks containing either large or small problem sizes. Behavioral results suggested that the performance did not differ between females and males. Neuroimaging data revealed that sex-related patterns of problem size effect were found in the conventional arithmetic circuits, including the left middle frontal gyrus (MFG), left intraparietal (IPS), and insula, with females demonstrating substantial brain responses of problem size effect compared to males. Moreover, the machine-learning method over the brain signal levels within the fronto-parietal circuits is discriminable of the sex/gender of human adults. These results demonstrated sex/gender effects in the activating patterns varying as a function of the distinct math problem complexity, even in a simple calculation task. Accordingly, our findings suggested that females and males use two complementary brain resources to achieve equivalently successful performance levels and highlight the pivotal role of neuroimaging facilities in uncovering neural mechanisms that may not be behaviorally salient.

英文關鍵詞： biological sex, mental arithmetic, mathematical cognition, fMRI, parietal cortex, prefrontal cortex

科技部補助專題研究計畫成果報告

（□期中進度報告/■期末報告）

以磁振造影研究探討數學能力之性別差異

計畫類別：■個別型計畫　□整合型計畫

計畫編號：110-2629-H-004 -001 -

執行期間： 2021 年 8 月 1 日至 2022 年 10 月 31 日

執行機構及系所：政治大學心理系

計畫主持人：張葶葶

共同主持人：范揚騰

計畫參與人員：陳乃鳳 范惠君 伍贊達 陳心喻 黃如君

本計畫除繳交成果報告外，另含下列出國報告，共 _1_ 份：

□執行國際合作與移地研究心得報告

■出席國際學術會議心得報告

期末報告處理方式：

1. 公開方式：

□非列管計畫亦不具下列情形，立即公開查詢

■涉及專利或其他智慧財產權，■一年□二年後可公開查詢

2.「本研究」是否已有嚴重損及公共利益之發現：■否 □是

3.「本報告」是否建議提供政府單位施政參考 ■否 □是，____（請列舉提供之單位；本部不經審議，依勾選逕予轉送）

中　華　民　國 111 年 12 月 30 日

目錄

中文摘要

一般大眾常認為數學能力有性別差異，且男性表現優於女性，然而實際文獻所報告之結果並不一致，且文獻間存在相當大之差異。根據TIMSS 1999、2003、2007及PISA 2006之成績統計，本國之四年級與八年級學生在科學與數學成績均無顯著之性別差異。過往雖不乏以大數據探討性別差異之行為與神經機制議題之研究，然直接檢驗數學認知下性別差異的磁振造影研究迄今為止僅有四篇文獻，且結果並不一致。為檢驗此議題，本研究以行為實驗及認知能力評估檢驗數學認知與學習之性別差異，並結合磁振造影實驗以檢驗數學認知與學習性別差異之神經機制，探討大腦結構與功能在進行算術作業時是否有性別差異，結果發現即便在行為表現上並無差異，女性卻在計算較複雜問題時前額葉及後頂葉活化比起男性強，使用機器學習模型得到此前額葉後頂葉之活化情形能夠預測，顯示兩性各自用各自的方式在進行相同的算術功能。本研究提供以教育、心理學、及認知神經科學領域探討男性及女性認知功能生理機制之實徵證據。

中文關鍵字： 生理性別、心算、數學認知、磁振造影、頂葉、前額葉

Abstract

Numerous empirical studies have reported that males and females perform equally well in mathematical achievement. However, still to date, very limited is understood about the neural mechanisms of whether and how men and women demonstrate differences when solving mathematical problems. The present study aimed to tackle this issue by manipulating arithmetic problem complexity and investigating functional significance using fMRI in young adults. Participants were instructed to complete two runs of simple calculation tasks containing either large or small problem sizes. Behavioral results suggested that the performance did not differ between females and males. Neuroimaging data revealed that sex-related patterns of problem size effect were found in the conventional arithmetic circuits, including the left middle frontal gyrus (MFG), left intraparietal (IPS), and insula, with females demonstrating substantial brain responses of problem size effect compared to males. Moreover, the machine-learning method over the brain signal levels within the fronto-parietal circuits is discriminable of the sex/gender of human adults. These results demonstrated sex/gender effects in the activating patterns varying as a function of the distinct math problem complexity, even in a simple calculation task. Accordingly, our findings suggested that females and males use two complementary brain resources to achieve equivalently successful performance levels and highlight the pivotal role of neuroimaging facilities in uncovering neural mechanisms that may not be behaviorally salient.

Keywords: biological sex, mental arithmetic, mathematical cognition, fMRI, parietal cortex, prefrontal cortex

(This report has been submitted to Journal of Neuroscience Research, and the revised version is under review)

1. Introduction

Over the past decades, empirical studies have reached the consensus that males and females perform equally well in arithmetic learning and mathematical achievement (Hyde, 2014). However, still to date, women showed less positive attitudes, lower motivation, and self-confidence toward mathematical learning than males (Rodríguez et al., 2020), and eventually remained minorities in personal choice of math-associated fields. This sex bias can deteriorate the male-math stereotype and continue to cause women's avoidance of mathematical learning. In searching for fundamental differences in the mechanisms of mathematical problem solving between males and females, numerous behavioral studies have extensively compared performance discrepancies between male and female students. However, understanding whether and how each sex/gender demonstrates specialty in the neural underpinnings is still very limited. In this study, we systematically investigate the distinctiveness of brain activations underlying mathematical problem solutions of each sex using the fMRI techniques. Given that it is difficult to discriminate whether differences between males and females are wired in the brain by nature or are learned from experience and environment, we adopted the term "sex/gender differences" (Chang et al., 2022; Jordan-Young & Rumiati, 2012; Springer et al., 2012) to capture both the biological mechanisms and the psychosocial expression of maleness and feminineness throughout the manuscript. By uncovering these issues, we seek to achieve more genuine sex/difference equality with more clarified investigations of the learning mechanism of individuals.

Numerous behavioral studies have extensively investigated sex/gender differences in school mathematical performance in recent decades (cf. Chang et al., 2022). Using large-scale meta-analytical analyses approach over millions of global participants, multiple studies have shown that sex/gender effects in mathematical performance, regardless of the contents, are subtle (Hyde et al., 1990; Hyde et al., 2008; Lindberg et al., 2010). Other studies have reported that sex/gender differences in mathematical performance, though negligible, declined with time. Studies and assessments administered and compared male and female students continued to show reduction in differences between 1973 and 2019 (Hyde et al., 1990; Mullis et al., 2020). Despite that some studies reported sex/gender differences at the individual level rather than reflecting societal group norms, for example, in the variance of performance (Baye & Monseur, 2016; Benbow et al., 2000; Lindberg et al., 2010) and specific problem types (Lindberg et al., 2010), the effects remained small and variable. Together these results have led researchers to agree that girls and boys reach parity in mathematical performance. Yet, understanding the underlying neural mechanisms of sex/gender effect in mathematical problem solving is still limited, as it not only captures contemporary differences in brain and behavior but also provides exclusive brain bases knowledge that is unseen in behavioral outcomes alone.

Neuroimaging studies have consistently identified distributed neural circuits activated during mathematical performance. As the vast majority of neuroimaging studies addressing math problem solving had emphasized calculation skills, here we focus on arithmetic problem solving. These neural circuits associated with arithmetic problem solution primarily encompass several nodes within the fronto-insular-parietal network, including anterior insula (AI), dorsal anterior cingulate cortex (dACC), dorsal posterior parietal cortex (PPC), and dorsolateral prefrontal cortex (DLPFC) (Arsalidou & Taylor, 2011; Chang et al., 2016; Houde et al., 2010; Menon et al., 2014; Ng et al., 2021). Within this set of networks, the IPS within the PPC is considered to play the most crucial role in representing and manipulating quantitative information (Ansari, 2008; Cohen Kadosh et al., 2008; Dehaene et al., 2003). Outside of the PPC, the canonical neural circuits include AI, dACC, and DLPFC (Cai et al., 2016; Chang et al., 2019; Levy & Wagner, 2011; Ng et al., 2021). The AI coupling with dACC forms the major components of the salience network (SN) (Menon, 2015b; Seeley et al., 2007) that is associated with subjective salience of external stimuli and in contributions to complex cognitive processes, including central executive function and affective processing. DLPFC, together with PPC, comprise the major nodes of the central executive network (CEN), engaged in information retention and manipulation during working memory, manipulation of quantities over epochs, construction of problem solutions, and decision making (Chang et al., 2019; Menon, 2015a; Miller & Cohen, 2001; Petrides, 2005; Rottschy et al., 2012). In a recent fMRI study, Chang and colleagues demonstrated that brain response profiles associated with judging sentences that required one-step arithmetic operations were associated with greater engagement and stronger within-network connectivity in this set of fronto-insular-parietal circuits relative to judgment over parallel narratives without any numerical information (Chang et al., 2019). Furthermore, the fronto-insular-parietal network has also been identified when assessing arithmetic problem-solving skills in the developmental progression across critical learning stages from early childhood to adulthood (Arsalidou & Taylor, 2011; Chang et al., 2019; Chang et al., 2016). Collectively, these results supported that the interconnected network jointly engages and synchronizes to form the network contributing to the core neural substrates of numerical problem-solving skills, ranging from simple number comparisons to complex arithmetic and problems that require mathematical reasoning and across the essential learning stage (Cho et al., 2012; Rosenberg-Lee et al., 2015; Rosenberg-Lee et al., 2011; Supekar & Menon, 2012). Yet, it remained unknown whether males and females showed distinctiveness in the set of the fronto-insular-parietal nodes, particularly during mathematical problem solving.

Numerical properties modulate the canonical arithmetic circuits, for example, problem complexity and problem size (Chang et al., 2016; Chang et al., 2015; De Smedt et al., 2011; Metcalfe et al., 2013; Stanescu-Cosson et al., 2000). The problem size effect refers to the problem complexity cost such that arithmetic problems with larger problem operands (e.g. 7+9; 6×8) were responded less accurately and slower than problems with smaller operands (e.g. 2+3; 2×4) (Campbell & Xue, 2001; De Smedt et al., 2011; Stanescu-Cosson et al., 2000). The effect of problem size likely reflects the specificity of strategy usage in distinct problem types. In particular, small problems are usually solved by fast-retrieving arithmetic knowledge facts, while large problems are solved by reasoning through the process of multistep calculations (Barrouillet et al., 2008; Campbell & Xue, 2001; De Smedt et al., 2011). Aside from the behavior findings, neural correlates of the problem size effect are also documented (De Smedt et al., 2011; Stanescu-Cosson et al., 2000). Stanescu-Cosson and colleagues demonstrated that large arithmetic problems engage more activations over the DLPFC and bilateral IPS. In contrast, small problems inversely engage stronger angular gyrus than large problems (Stanescu-Cosson et al., 2000). Several other studies have also reported similar results with school-age children (Chang et al., 2016; Chang et al., 2015), with the exception that in children, it is hippocampus rather than the angular gyrus (AG) shows stronger activations for small problems (Cho et al., 2012; De Smedt et al., 2011). In sum, the results of these studies provided additional biological support for the involvement of procedure-based computation and working memory allocation when solving complex problems with larger sizes, as well as retrieval of mathematical facts when solving simple problems. Given that problem size effect is consistent and reliable across studies to probe arithmetic-associated brain responses, it is likely suited to address important questions of differences in the neural mechanisms between sexes/genders.

Even until today, only four studies have used fMRI techniques to directly compare brain processing of males and females during arithmetic problem solving. Wang et al. (2007) compared brain responses of males and females during a high-pressured serial subtraction of 13 from a 4-digit number and a low-pressure backward counting from 1000. They found that the right PFC was more active in males than in females while performing the stressed task. Subsequently, Keller and Menon (2009) compared sex/gender differences in the brain activations while participants calculated 3-operand addition and subtraction problems. The results suggested that males engaged in a greater level of IPS, AG, lingual and parahippocampal gyri, whereas no regions showed greater functional activation in females than males. Paradoxically, a reverse pattern of sex/gender differences was found in the voxel-based morphometry with the subset of the samples. Females showed greater volume and density than males in the regions that were activated by the arithmetic task. Pletzer (2016) examined the brain response patterns of young adults as they performed subtraction and multiplication tasks. In that study, participants showed stronger IPS activations for subtraction as well as greater AG activations for multiplication tasks. Interestingly, this operation effect was only observed in males but not in females, suggesting that females showed less differentiation between numerical problems of distinct nature. In a more recent study, Kersey and colleagues quantified brain responses while school-age children watched education videos depicting mathematics. To obtain the index of neural similarity, intersubject correlations were computed across all children's brain responses. According to the authors, intrasex and intersex neural similarity did not present differences in processing between boys and girls, including bilateral IPS, bilateral inferior frontal gyrus (IFG), and anterior cingulate cortex (Kersey et al., 2019). These results led the authors to conclude that there is much more similarity than differences between male and female brains in nature. Taken together, these previous attempts present a contradictory picture of neural dissociation between males and females associated with mathematical cognition. All these previous studies had varied in task designs, analysis strategies, and the sampling ages, making drawing specific conclusions about the sex/gender differences in mathematical cognition challenging. Nevertheless, it is worth noting that although functional responses differed between males and females, no compelling differences in behavioral measures were observed in the studies mentioned above.

In the current study, we attempt to systematically examine sex effect on brain response profiles during mathematical problem solution by collecting fMRI data from adults who were proficient in general arithmetic problem solving skills. In order to linearly control the task complexity with the corresponding problem solving strategies, we directly manipulate problem size as large and small problems since the arithmetic associated neural circuit, i.e. the fronto-insular-parietal network, has been consistently identified as a function of the problem size as reviewed above. We also applied a machine learning logistic regression model to assess whether the brain responses of the arithmetic task can discriminate between males and females. On the basis of existing mathematics-related assessment reports, there was a very subtle sex disparity between sexes. Therefore, we predicted that behavior performances would not show differences in this simple task. As previous literature reported that males and females can adopt distinct problem solving strategies, with boys tend to solve mathematical problems using fast rote-fact retrieving, estimation and insight strategies, whereas females tend to adopt more concrete, algorithmic calculation (Bailey et al., 2012; Gallagher et al., 2000; Zhu, 2007). We hypothesized that sex/gender differences will be observed in the wired mathematical learning-associated brain circuits. More specifically, we expected that females would show greater fronto-parietal engagement during the vital mathematical task giving their problem solving strategies, and machine learning methods over the fronto-parietal circuits would predict the sex/gender labels.

2. Method

2.1. Participants

Seventy-five adults (38 females and 37 males) were recruited from local educational institutions in Taipei city, Taiwan. Among the participating adults, four had excessive head movements (for the movement exclusion criteria, see the fMRI data preprocessing section below), resulting in the final sample of seventy-one participants (36 females; age range 18.93 to 29.09 years, M = 23.04, SE = 0.28).This sample size is adequate for the suggested number of at least 5 to 9 events per independent variable (EPV) by Vittinghoff and McCulloch (2007) for further logistic regression analysis. Mean ages did not differ between females (M = 22.60, SE = 0.37) and males (M = 23.49, SE = 0.41) ($t_{(69)}$ = 1.62, p = 0.11, 95% CI = [-1.99, 0.21], d = 0.38). All participants were right-handed with no reported history of psychiatric or neurological disorders and had normal or corrected-to-normal vision. All participants had comparable educational status (undergraduate or graduate students). 56 of the participants (28 females) completed an arithmetic assessment prior to the fMRI scan using an arithmetic test similar to the French Kit (Ekstrom, 1976) and our previous study (Chang et al., 2018). During the test, participants were instructed to solve a mixture of single- and two-digit addition and subtraction problems as quickly and accurately as possible. Mean accuracy of this test did not differ between females ($M = 0.730$, $SE = 0.022$) and males ($M = 0.755$, $SE = 0.027$) ($t_{(54)} = 0.703$, $p = 0.485$, 95% CI = [-0.094 0.045], $d = 0.188$). Informed written consent was obtained from each participant. All participants were volunteers and were treated according to the Helsinki Declaration guidelines. All study protocols were approved by the National Chengchi University Review Board.

2.2. Experimental design

All participants were instructed to complete two runs of 2-operand mathematical verification tasks during fMRI scanning. The problems of this verification task consisted of combinations of single-digit operands from 2 to 9, with the exclusion of tie problems (e.g., 5+5). Within each run, the stimuli included 56 addition and subtraction problems. Problems consist of two conditions: large- and small-sized problems. For addition, in the large problem condition, the product of the two operands was larger than 25 (e.g., 8+7); in the small problem condition, the product of the two operands was smaller than or equal to 25 (e.g., 5+2). For subtraction, stimuli were inverses of addition problems. Each trial began with a '*' sign as a fixation for 500 ms followed by the presentation of a problem for 3000 ms in the center. Next, the corresponded answer to the problem was displayed for 1000 ms. During this period, participants were asked to determine the correctness of the showing answer by pressing one of two keys based on their answer; 50% of the trials were correct (e.g., '6 + 3 = 9'), and the other 50% were incorrect (e.g., '6 + 3 = 8'). The incorrect answers differed by ± 1 or ± 2 of the correct ones. The screen was then blank for 750 ms. Afterward, the screen remained blank for a jittered inter-trial interval between 2 and 5 s. Each of the two runs lasted approximately 8 min. The presenting orders of the trial according to the problem size were randomized, and the performance sequence of the two runs was counterbalanced between the participants.

2.3. fMRI data acquisition

Neuroimaging data was acquired using a Siemens MAGNETOM Skyra 3 T scanner at National Chengchi University in Taipei City, Taiwan. Head movement was minimized during the scan using cushions placed around the head of each participant. T2* weighted echo-planar sequences were employed with the following parameters: TR = 2 s, TE = 30 ms, flip angle = 90°, 36 ascending axial slices with slice thickness = 4mm, field of view = 220 × 220 mm2, matrix size = 64 × 64, providing an in-plane spatial resolution of 3.4 mm. In the same scan session, high-resolution T1-weighted MRI sequences were acquired for each participant to aid localization of functional data, with the following parameters: TR = 3500 ms; TE = 3.37 ms; TI = 1100ms, flip angle = 7°, field of view = 256 × 256 mm2, matrix size = 256 × 256, resulting in resolution of 1 ×1 ×1 mm3, number of excitations = 1, 192 slices in axial plane.

2.4. fMRI data preprocessing

SPM12 (http://www.fil.io.ucl.ac.hk/spm) was used for preprocessing of functional MRI data. All functional images were corrected prior to statistical analysis for errors in slice timing, realigned to the first image of each run to correct for head motion, coregistered to each of the individual participant's structural scans, normalized to standard stereotaxic space (based on the Montreal Neurologic Institute coordinate system), and smoothed with a 6 mm full-width half-maximum Gaussian kernel to decrease spatial noise. Participants with movement more than 3 mm in translational directions and 3 degrees in rotational directions were excluded from further analyses. The average movements of the final participants were 0.34 (SE = 0.01), 0.47 (SE = 0.03), and 0.95 (SE = 0.05) mm in the x, y, and z directions, with 0.84 (SE = 0.05), 0.34 (SE = 0.02), and 0.28 (SE = 0.01) degrees of roll, pitch, and yaw, respectively.

2.5. Individual and group-level analyses

Statistical analysis was performed on both individual and group-level data using the general linear model (GLM) implemented in SPM12. Individual subject analyses were first performed by applying GLM that modeled the correctly responded trials as regressors and convolved with a canonical hemodynamic response function to model the expected BOLD signal. Incorrectly responded trials, the epoch participants made responses, and the six motion parameters generated in the SPM12 realignment procedure were included as regressors of no interest. Voxel-wise t-maps for each effect of interest from individual level were entered into a random-effects 2 (problem size) × 2 (sex) mixed-design ANOVA, with problem size as within-subject factors and sex as between-subject factor. We investigated the main effects and interactions at the brain level. Because F-tests do not test the direction of the effects, t-contrasts were calculated for visualization in the subsequent analyses to determine the direction of any significant effects. All significant results were determined according to a voxel-wise height threshold of $p < .005$ uncorrected, and a multiple comparison correction at a spatial-extent threshold of FWE $p < .05$ after gray matter masking.

2.6. Logistic regression and cross-validation

Logistic regression analysis was implemented to estimate whether the brain response that showed a problem size effect could predict sex/gender group labels. Because neural circuits associated with arithmetic problems predominantly include fronto-insular-parietal regions (Arsalidou & Taylor, 2011; Chang et al., 2019; De Smedt et al., 2011; Houde et al., 2010), we conducted logistic regression and cross-validations within these circuits using ROI (region of interests) approach. In order to avoid inflated correlations produced by deriving ROIs from the same dataset (Vul et al., 2009), we defined ROIs using a meta-analysis based on the approach of our previous work (Chang et al., 2019; Chang et al., 2018). Specifically, a Bayesian meta-analysis of the reverse inference mask available in Neurosynth (Yarkoni et al., 2011) was conducted using the search term ‘arithmetic’, resulting a total of 96 studies generated. A false discovery rate (FDR) adjusted *p* value of 0.01 was applied to produce the association test map. The coordinates with peak z-scores with clusters exceeding 50 voxels on the association test map were identified using the xjView toolbox (www.alivelearn.net/xjview), and selected for further analyses. The resulting brain maps encompassed the left IPS (peak at [-28, -60, 44]), the right IPS [30, -64, 46], the left insula [-22, 22, 2], the left MFG [-26, 10, 54], the left IFG [-50, 10, 26], and the left superior frontal gyrus (SFG) [-4, 14, 54]. In subsequent logistic regression analyses, a 10-mm radius sphere (515 voxels with voxel size = 4120 mm^3) centered on each of these six identified peak coordinates was created using MarsBaR (http://marsbar.sourceforge.net/) as selected ROIs. Estimated beta values of activation level differences between large and small problems extracted from these ROIs were then entered into the following logistic regression model to classify participants based on their sex/gender.

A multiple logistic regression model was built and verified using the forward stepwise method based on the Akaike information criterion (AIC) (Akaike, 1974) selection and the probability of the Wald statistic. The AIC measures the trade-off between the uncertainty in a model and the number of predictor variables in the model. Lower AIC values imply better prediction of sex/gender labels, as they explain the greatest amount of variation in the response variable with the least amount of predictor variables. The forward stepwise logistic regression starts with a null model, adds the most contributed variables one by one, and ends with a model that picks the best variables for an optimal solution. In the current study, the beta value differences between large and small problems generated from the 6 Neurosynth ROIs were considered as predictive variables. The optimal subset of variables related to sex label discrimination could be determined by utilizing the forward stepwise selection method.

Finally, the classification accuracy was evaluated using a stratified k-fold cross-validation procedure. This evaluation procedure consists of three steps. In the first step, the samples were randomly partitioned into a training set (70%) and a test set (30%), both sets contain approximately the same percentage of samples of each target class (females and males approximately 1 / 1) as the total number of participants (females: males = 36: 35). The test set was held for later estimation of the generalizability of the model classifier. In the second step, the training set was further shuffled and divided into k (k = 10) equal-sized folds. One fold was used for performance validation and the remaining k-1 folds were combined into a sub-training set for model fitting. Again, a similar proportion of samples (females and males) are included in each set as in the total number of participants. After that, the above k-fold cross-validation procedures were repeated ten times on the given sub-training dataset. We then obtained the average classification accuracy of training set in the second step. In the last step, we estimated the test set defined in the first step, computing the classification accuracy (proportion of all participants’ sex correctly predicted), sensitivity, and specificity. The receiver-operating characteristic (ROC) curve was constructed using the probability thresholds with corresponding data points (sensitivity, 1 - specificity), and the area under the curve (AUC) was then calculated.

The modeling and statistical analyses were implemented using R packages 'caret', and 'MLeval'.

2.7. Voxel-based morphometry (VBM) analysis

Sex-related differences in brain anatomy were examined using VBM analysis. The CAT12 toolbox (CAT12; http://dbm.neuro.uni–jena.de/cat12) implemented in the SPM12 software was used to process the T1-weighted images. Structural T1-weighted images of each participant were first converted into the Neuroimaging Informatics Technology Initiative (NIFTI) format through SPM12. The images were then pre-processed with the standard default procedure recommended in the CAT12 manual. The preprocessing steps included skull stripping, segmentation into gray matter (GM), white matter (WM), and cerebrospinal fluid (CSF), followed by spatial normalization to the DARTEL template in the Montreal Neurological Institute (MNI) space with 1.5 mm cubic resolution. The quality of the images was assessed with the built-in image quality rating and manually visual check. Finally, images were smoothed using a 6-mm full-width half-maximum (FWHM) isotropic Gaussian kernel. In addition, the total intracranial volume (TIV), which is the sum of GM, WM, and CSF volumes in the native space, was also estimated.

3. Results

3.1. Behavioral results

The mean accuracy and reaction time for each problem condition for each participant were computed and analyzed using repeated-measures ANOVA with problem size (small, large) as within-subject factors and sex (female, male) as a between-subjects factor. For the accuracy (Figure 1A), as predicted, there was a significant main effect of problem size, showing that participants performed more accurately on small size problems than large size problems (96.8% vs. 95.6%, $F(1,69) = 9.334$, MSE = 0.005, $p = 0.003$, $\eta^2 = 0.016$). No differences between males and females was observed when performing the arithmetic task (95.4% vs. 96.9%, $F(1,69) = 1.931$, MSE = 0.008, $p = 0.169$, $\eta^2 = 0.024$), nor did the interaction effect between sex and problem size was significant ($F(1,69) = 0.618$, MSE < 0.001, $p = 0.434$, $\eta^2 = 0.001$) , indicating females and males perform equally well on this current simple single-digit calculation task.

Regarding the reaction time analysis (Figure 1B), likewise, there were main effects of problem size, showing that participants responded faster to small size problems than to large size problems (640 ms vs. 650 ms, $F(1,69) = 9.969$, MSE = 3901, $p = 0.002$, $\eta^2 = 0.002$). No difference was observed between males and females (637ms vs. 652ms, $F(1,69) = 0.227$, MSE = 7982, $p = 0.635$, $\eta^2 = 0.003$), nor did the interaction effect between sex and problem size was significant ($F(1,69) = 0.244$, MSE = 96, $p = 0.623$, $\eta^2 = 0.001$).

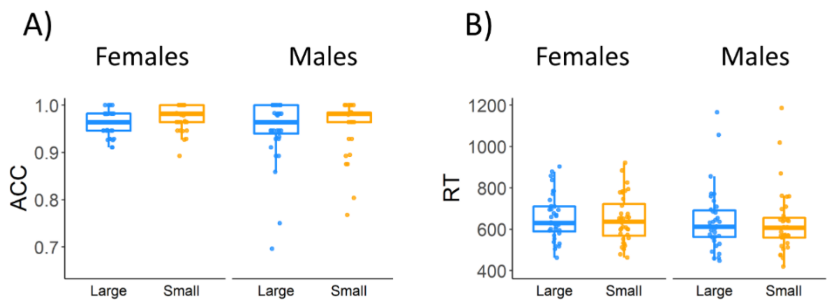


Figure1. Accuracy (ACC) and Reaction times (RT) of the arithmetic task in females and males. (A) Participants performed more accurately to small size problems than large size problems. No significant difference between sexes was found, and sex and problem size did not interact significantly. (B) Participants responded faster to small size problems than large size problems. No significant difference between sexes was found, nor was the interaction effect between sex and problem size significant.

3.2. Brain imaging results

3.2.1. Brain responses that showed differences between large and small problems

We first identified brain regions showing response differences associated with arithmetic problem size by contrasting the neural correlates of large and small problems in the pooled group of males and females. The problem size effects on brain activation are presented in Figure 2. First, across all participants, relative to small problems, large problems exhibited a widespread fronto-parietal network of regions, including bilateral MFG extending to adjacent IFG and medial frontal gyrus in the prefrontal cortex (PFC), bilateral IPS in the PPC. Additional clusters were also found in the ventrotemporal occipital cortex (VTOC), including bilateral lingual gyrus (LG), fusiform gyrus (FG), and calcarine. In contrast to large problems, small problems were activated more in the bilateral supramarginal gyrus (SMG) and the left AG in the PPC, medial prefrontal cortex, and posterior cingulate cortex, and bilateral superior temporal gyrus (STG) (Table 1 shows the detailed results of the peak coordinates in each cluster). Simple main effect results suggested that the problem size effect is more salient in females, whereas males yielded a less profound pattern across these distributed regions (Figure 2 and Table 2).

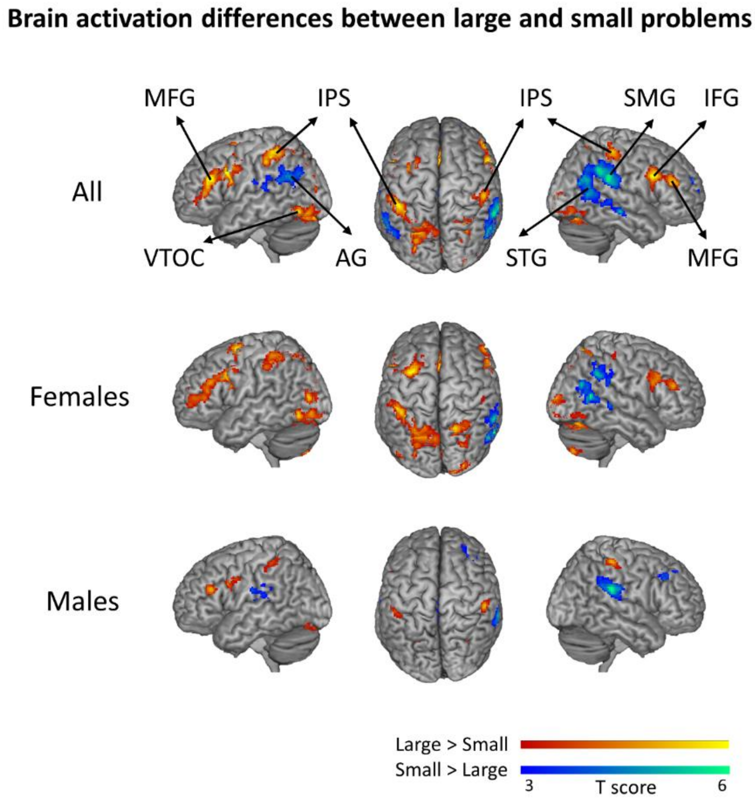


Figure 2. Brain regions that showed different activation levels between large- and small-size problem solving across overall participants (upper panel), in females (middle panel), and in males (lower panel). Activations in fronto-parietal regions including MFG (middle frontal gyrus), IFG (inferior frontal gyrus) and IPS (intraparietal sulcus) were greater in large problems compared to small problems. On the other hand, activations in AG (angular gyrus), SMG (supramarginal gyrus) and STG (superior temporal gyrus) were higher in small problems than large problems.

Table 2 Sex differences of the problem size effect

Region	Corrected p_{FWE}	# of voxels	Peak T-score	Peak MNI coordinates		
				x	y	z
Problem size effects (females)						
Large > Small						
R lingual gyrus	<0.001	13987	9.48	20	-82	-6
L inferior frontal gyrus	<0.001	3847	6.55	-38	2	34
R intraparietal sulcus	<0.001	1948	6.69	34	-48	44
L anterior cingulate gyrus	<0.001	1583	6.47	-8	18	46
R middle frontal gyrus	<0.001	710	4.94	50	38	22
Small > Large						

R superior temporal gyrus	0.017	378	5.28	62	-56	10
R medial frontal gyrus	0.022	359	4.04	4	48	-8
R supramarginal gyrus	0.032	333	4.81	62	-48	36
Problem size effects (males)						
Large > Small						
L middle frontal gyrus	<0.001	894	5.25	-48	30	24
L precuneus	<0.001	836	4.94	-26	-68	38
R lingual gyrus	<0.001	765	7.86	16	-82	-10
L lingual gyrus	0.001	596	5.85	-14	-88	-10
R intraparietal sulcus	0.001	566	5.16	52	-30	54
Small > Large						
R supramarginal gyrus	<0.001	936	6.27	58	-32	24
R posterior cingulate cortex	<0.001	864	4.81	8	-32	44
L supramarginal gyrus	0.001	530	4.81	-50	-26	14
R middle frontal gyrus	0.011	375	4.61	40	26	38

3.2.2. Females exhibited larger problem size effects than males

To investigate whether male and female show differences when processing large and small problems, we examined brain areas that showed problem size by sex interaction. This analysis revealed significant differences in the left MFG, IPS, and the right dACC (Figure 3; Table 3 reveals detailed results of the peak coordinates in each cluster). Further analysis of the averaged beta weights of each significant cluster revealed that the interaction effect was driven by the problem complexity cost (Complex-Simple) being more prominent in females (Figure 3). Specifically, females showed stronger activations for complex then simple problems in the left MFG ($t_{(35)} = 4.006$, $p < 0.001$, 95% CI = [0.120, 0.366], $d = 0.677$), the left IPS ($t_{(35)} = 6.600$, $p < 0.001$, 95% CI = [0.371, 0.701], $d = 1.116$), and the right dACC ($t_{(35)} = 4.332$, $p < 0.001$, 95% CI = [0.098, 0.272], $d = 0.732$). Males, on the contrary, showed a minimal or null effect of problem complexity in these brain regions (left MFG ($t_{(34)} = -2.033$, $p = 0.05$, 95% CI = [-0.228, -0.001], $d = 0.349$); the left IPS ($t_{(34)} = 1.699$, $p = 0.099$, 95% CI = [-0.022, 0.246], $d = 0.291$), and the right dACC($t_{(34)} = -1.713$, $p = 0.096$, 95% CI = [-0.124, 0.011], $d = 0.294$)).

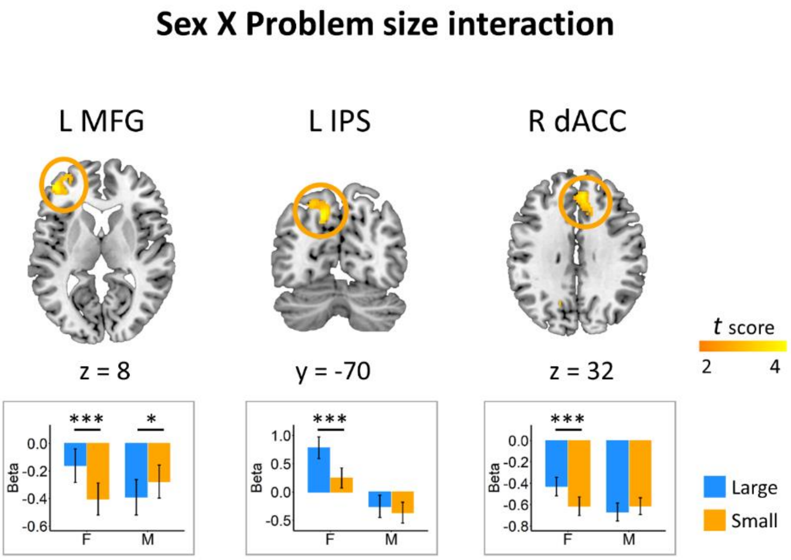


Figure 3. Statistical maps illustrating regions activated for sex and problem size interaction effects. A problem size effect was evident in females (F), with greater activation in large problems than small problems, whereas a problem size effect was negligible in males (M). Error bars represent standard errors. **p* < .05, ****p* < .001. Abbreviations: L MFG (left middle frontal gyrus), L IPS (left intraparietal sulcus), R dACC (right dorsal anterior cingulate cortex).

Table 3 Sex differences in brain activation in the mental arithmetic task

Region	Corrected p_{FWE}	# of voxels	Peak T-score	Peak MNI coordinates		
				x	y	z
Problem size effects						
Females > Males						
R dorsal anterior cingulate gyrus	0.006	447	4.06	6	40	32
L intraparietal sulcus	0.006	444	4.33	-14	-70	40
L middle frontal gyrus	0.039	314	4.29	-44	48	8
Males > Females						
No significant clusters						

R, right. L, left.

3.2.3. Brain responses in fronto-parietal circuits predict sex/gender difference

We then examined whether brain activity that showed problem size effects could accurately distinguish females and males using a logistic regression function. The averaged beta values of large and small problem size for each participant were extracted from the six ROIs defined by meta-analysis to avoid inflated correlations, as introduced in section 2.3.4. The selected ROIs were highly overlapped with the activation level maps generated from the one-sample t-test on the contrast of Complex minus Simple problems on the data from pooled males and females together (Figure 4). In order to investigate whether the regional brain response profile could predict the sex/gender label, we conducted a binary logistic regression classifier to categorize males and females using the estimated activation level difference between large and small problems extracted from the aforementioned unbiased ROIs. The results showed that among the six ROIs, the logistic coefficients were significant in the left insula (beta(β) = 6.089, p = 0.011, odds ratio = 441.099, 95% CI = [1.708, 11.167]), the left MFG (beta(β) = 2.677, p = 0.006, odds ratio = 14.541, 95% CI = [0.918, 4.761]), and the left IPS (beta(β) = 2.094, p = 0.005, odds ratio = 8.116, 95% CI = [0.723, 3.688]) (Figure 4). These results indicated that female participants had a higher probability of exhibiting greater degrees of activation level difference between large and small problems within the left insula, MFG, as well as IPS (Figure 5).

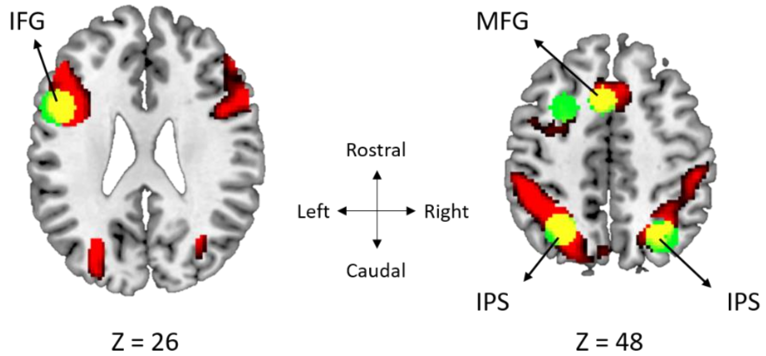


Figure 4. Brain regions that showed overlapping between the whole brain analysis of problem size effect and the selected ROIs based on meta-analysis results. The simple main effect of problem size (red), selected ROIs (green), and regions of overlap (yellow) on the standard space. Coordinates are in MNI space (mm).

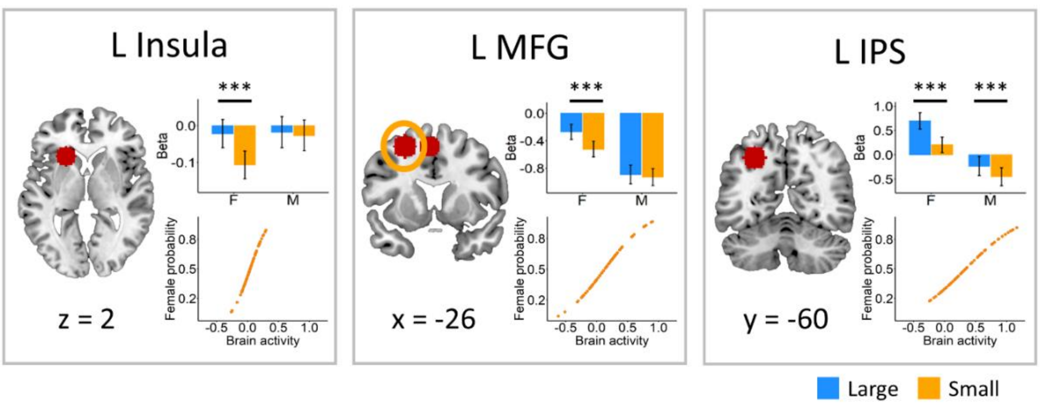


Figure 5. Logistic regression results of ROIs that brain activity-based classification successfully classified participants' sex. Within each frame, the top-right bar plots revealed that

problem size effects were only observed in females (F) over left insula and left middle frontal gyrus (L MFG) while the effects were found in both females and males (M) over left intraparietal sulcus (L IPS). The bottom-right sigmoid function plots of each frame indicated that participants' sex (y axis) could be classifiable based on brain activation level differences between large and small problems (x axis).

In an attempt to build a logistic model that best describes the sex differences in problem complexity, the stepwise forward method was then performed with the variables based on the estimated activation level difference between large and small problems extracted from the selected ROIs. Of the six aforementioned ROIs, the optimal subset of variables related to sex label discrimination using the forward stepwise selection method resulted in five ROIs – right IPS, left IPS, left IFG, left MFG, and left insula. The AIC of the final model consisting of the five ROIs is 77.467, with the logistic coefficients for right IPS (beta(β) = -3.682, Wald χ^2 = 10.4), left IPS (beta(β) = 6.228, Wald χ^2 = 8.7), left MFG (beta(β) = 3.482, Wald χ^2 = 4.9), left IFG (beta(β) = -2.537, Wald χ^2 = 4.2) showed statistically significant at the 0.05 level (Figure 6A). All estimation parameters are reported in Table 4. The overall model classification accuracy was then evaluated using a three-step cross-validation procedure, summarized in Figure 6. Figure 6B illustrated the machine learning method's analyzing strategy and the model evaluation procedures. First, the entire data set of the five-variant logistic regression model was randomly split into the training subset (70%) and the testing subset (30%). Second, a stratified 10-fold repeated cross-validation procedure was computed on the training data, resulting in an average accuracy of 71% (kappa = 0.41). Figure 5C shows the distribution of the performance measure reported by stratified 10-fold repeated cross-validation. Finally, the resulting model was evaluated in the remaining test subset. The classification accuracy of multiple logistic regression model is 81% (95% CI = [58% - 95%], kappa = 0.63), the sensitivity/specificity of the model is 0.67/1, resulting in an AUC of 0.91. The ROC curve and the AUC derived from the test set were summarized in Figure 6D.

Table 4 Multiple logistic regression model results for predicting biological sex during simple and complex arithmetic task

term	β	S.E.	Wald χ^2 (df = 1)	*p*. value	OR	95% CI OR
Intercept	-0.69	0.45	2.3	0.128	0.50	-1.61 – 0.18
R IPS	-3.68	1.14	10.4	0.012	0.03	-6.20 – -1.67
L IPS	6.23	2.11	8.7	0.003	506.68	2.50 – 10.88
L MFG	3.48	1.58	4.9	0.027	32.53	0.68 – 6.94
L IFG	-2.54	1.24	4.2	0.040	0.08	-5.13 – 0.21
L Insula	4.88	3.52	1.9	0.165	132.22	-1.75 – 12.32

R, right. L, left. IPS, intraparietal sulcus. MFG, middle frontal gyrus. IFG, inferior frontal gyrus. Significant variables are in bold.

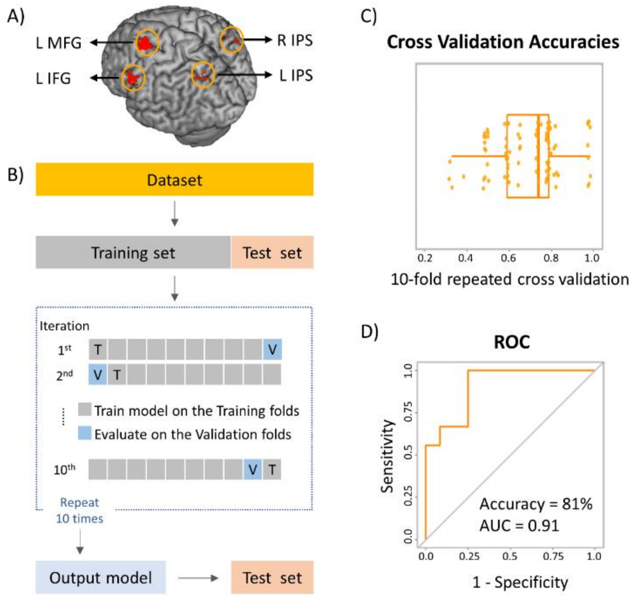


Figure 6. Multiple logistic regression results. (A) Final multiple logistic model explaining sex differences in problem complexity by looking at the estimated activation levels between large and small problems extracted from the selected ROIs. In the final five-variant model, predictors of right IPS, left IPS, left MFG, and left IFG contributed significantly to female/male differentiation during the arithmetic task. (B) Flow-chart of multiple logistic regression model evaluation. In the original dataset, the prediction variables were beta value differences in five ROIs: right IPS, left IPS, left IFG, left MFG, and left insula. First, we shuffled and split up the original dataset of the multiple logistic regression models into training (70%) and test (30%) subsets. Nest, we performed a 10-fold repeated cross-validation procedure on the training data to generate an average output model performance over the repeated 10 folds. Last, the remaining test data was used to evaluate the output model. (C) The classification accuracy performances of the training set were reported by 10-fold repeated cross-validation. (D) Receiver operating characteristic curve (ROC) and overall model performances for prediction of sex labels on the test set. AUC indicates area under the curve.

3.2.4. Sex/gender effect on neuroanatomical structure

Finally, we investigated whether sex/gender differences in functional brain activations were elicited from changes in fundamental neuroanatomical differences. We focused on gray matter volume in the three fronto-parietal regions identified in the functional activation analysis which showed sex related problem size effect – dACC, left IPS, and left MFG. After controlling for the total intracranial volumes, all three regions showed sex-related differences in regional volume, with males showed larger cortical volume in the above regions (Figure 7) (dACC, $p = 0.02$; left IPS, $p \leq 0.001$; left MFG, $p = 0.01$). These results showed that the increase in female fronto-parietal activation is not related to changes in the underlying neuroanatomy.

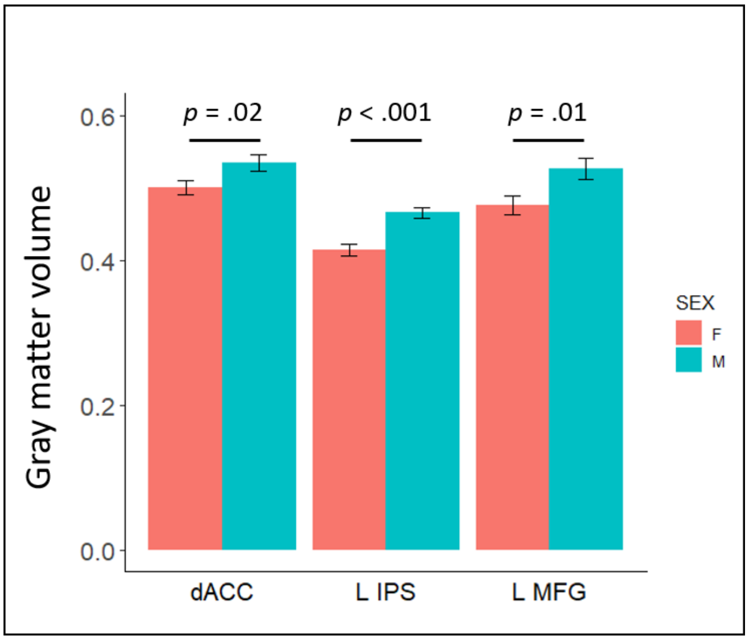


Figure 7. Regional volumes in the functional clusters within the fronto-parietal regions showed sex-related activation differences. Sex-related differences in gray matter volume were observed in the dACC, left IPS, and left MFG. Male participants showed larger cortical volume in these three regions.

4. Discussion

In this study, we investigate whether the problem complexity of the arithmetic task modulates brain responses differently in females and males. We directly manipulated problem complexity by varying problem size, aiming to reveal the activity profiles of crucial math processing in both sexes. As far as we are aware, our findings are the first to examine the problem size effect in each sex/gender, as it has strong potential to represent the effectiveness of strategies used by each individual (Cho et al., 2011). As predicted, we did not observe any behavioral performance differences between females and males. However, the sex-/gender-related effects on neural responses varied depending on problem complexity. This interaction is manifested by females showing greater fronto-parietal activation for complex problems than males. More specifically, sex effects on problem size were observed in left MFG, IPS, and right dACC, with females exhibiting greater activations in large problems than in small problems. Crucially, the machine learning algorithm revealed that the fronto-parietal signal levels during arithmetic tasks could successfully discriminate males from female participants. These findings collectively suggest that the brain responses while performing mathematical tasks are different in each sex, particularly in the fronto-parietal circuits.

4.1 Males and females showed similar behavior performance

In line with earlier behavior assessment (Hyde, 2014) and task-dependent neuroimaging studies (Keller & Menon, 2009; Pletzer, 2016), our results showed no differences between females and males in either accuracy or reaction times. Note that we implemented the presentation duration of the stimuli long as 3-second not only to ensure participants had enough time to obtain each problem solution but also to avoid motor responses contaminating the neural responses toward responding to numerical problems, as the supplementary motor area (SMA) is consistently activated during arithmetic problem solving (Menon et al., 2014). Participants were instructed to make a verification response immediately after the problem offset. Therefore, the behavioral results may not be valid for indexing the actual time to respond. Nevertheless, our results still inherent conventional problem size effect in the measurements of accuracy, response latencies, and brain response profiles (De Smedt et al., 2011; Stanescu-Cosson et al., 2000), suggesting that the task design has sufficient loading to differentiate the processes between distinct conditions even when the performance reaches high as ceilings. The current task design is thus sensitive enough to provide behavior-independent evidence of examination of brain functional organization.

4.2 Problem size effect in fronto-parietal circuits is more salient in females

The key finding of the current study is that the brain functions differently to problem complexity between females and males in the left MFG, left IPS, and dACC. Within these regions, females exhibited robust problem size effects, whereas males displayed negligible effects. When managing math problem solving, these three regions are activated as part of the fronto-parietal arithmetic circuits. Functional imaging studies have identified the contributions of these nodes to mathematical cognition. For instance, IPS has been identified as playing a crucial role in quantity representation (Arsalidou & Taylor, 2011; Dehaene et al., 2003) and has been suggested to reflect the use of quantity-based procedure strategies in mathematics (Stanescu-Cosson et al., 2000). Brain activations are typically increased for large compared to small problems in IPS (De Smedt et al., 2011; Polspoel et al., 2019; Tiberghien et al., 2019). On the other hand, MFG has been associated with complex and effortful tasks involving quantity manipulation (Chang et al., 2015; Menon et al., 2000; Wu et al., 2009). Additionally, ACC coupling with insula constitutes the salience network, which serves as a major causal hub in complex problem solving, functions as integrating and directing salient stimuli and initiating control signals (Menon, 2015b). Overall, these findings suggested that females recruited greater neural resources than males during mathematical problem solving, even when solving simple math problems.

As suspected, one possibility of the sex/gender differences in brain response profiles can be attributed to the distinct problem solving strategies used by each sex/gender (Bailey et al., 2012; Gallagher et al., 2000; Quinn & Spencer, 2001; Zhu, 2007). Problems solved by procedural strategies are usually associated with strong activations within the fronto-insular-parietal circuits, including the bilateral IPS, MFG, insula, and ACC (Grabner et al., 2009; Sokolowski et al., 2022).

Such finding is in accordance with several behavioral reports of sex/gender differences in arithmetic problem solving strategies. Bailey and colleagues found that across the entire elementary school stage, boys tend to solve simple addition problems more often than girls (Bailey et al., 2012; Carr & Davis, 2001). Carr and colleagues found that girls retrieve less, but used more manipulative strategies than boys (Carr & Davis, 2001). Bailey interpreted the sex/gender difference in problem solving strategy as a product of personalities. In particular, males tend to be more competitive and risk-taking, whereas females are more risk-averse (Azanova et al., 2021; Bailey et al., 2012). In support of this claim, Quinn and Spencer (2001) found that females are more likely to choose an error-avoiding strategy rather than a speedy manner. Our results are in line with the previous observations and provide the biological bases of the sex/gender differences in the problem solving strategies.

Another possibility can attribute sex/gender differences in the attitude toward mathematics (Di Martino & Zan, 2011). Negative emotional reactions – the so-called math anxiety – can be elicited when dealing with math-related situations (Ashcraft & Ridley, 2005). Even when solving simple arithmetic problems, it can be triggered, especially during timed conditions (Caviola et al., 2017) and when tasks increase in complexity (Ashcraft & Krause, 2007). Negative correlations between math anxiety and math achievement have also been reported in a wide range of students (Hembree, 1990). Even though girls generally perform equivalently well with boys in mathematical achievement, females are notoriously high in self-report math anxiety (Devine et al., 2012; Else-Quest et al., 2010; Ferguson et al., 2015; Hembree, 1990; Lau et al., 2022; Maloney et al., 2012) that can likely be attributed to social-cultural or emotional factors (Beilock et al., 2007; Bieg et al., 2015). The sex-specific math anxiety profile remained even when general anxiety was controlled (Devine et al., 2012; Goetz et al., 2013). Math anxiety is also often comorbid with limited working memory (Ashcraft & Kirk, 2001; Ashcraft & Krause, 2007; Ramirez et al., 2013). One empirical example provided by Ashcraft and Kirk (2001) reported that highly math-anxious college students were less accurate in performing addition problems with carry operation only when implementing a secondary task that required a high working memory load. Consequently, the impact of math anxiety on learning can possibly be due to the disturbance of working memory strategies while performing mathematical tasks. Consistently, highly math-anxious participants showed more enhanced engagement of the fronto-parietal cortices, including the IPS and MFG (Supekar et al., 2015). Therefore, it is possibly the exceedingly high math anxiety that up-regulates fronto-parietal engagement in females during the timed calculation task. This interpretation, however, is admittedly speculative. To confirm the hypothesis, further direct assessments on the relationship between math anxiety level and brain response profiles of each sex/gender are still needed.

4.3 discrepancies with previous studies

Our current results contradict other fMRI studies that have probed the larger task effect on males rather than females (Keller & Menon, 2009; Pletzer, 2016). The discrepancies are most likely resulted from the varied task design and sampling variance. In the study conducted by Keller and Menon (Keller & Menon, 2009), a 3-operand mixed operation task was implemented and compared with a number identification task, resulting in greater dorsal and ventral-stream activations in males. The multi-step and multi-operation calculation task can consume more working memory load and require multi-strategy engagement, making it difficult to disentangle sex effect resulted from problem complexity or problem operation, In another study, Pletzer (Pletzer, 2016) compared two-digit subtraction with single-digit multiplication, demonstrating a dissociated activation map in males. Given that decomposition and transformation strategies are frequently reported in solving multi-digit subtraction with borrowing (LeFevre et al., 2006) , it may be necessary to use a combination of strategies and to engage higher order of attention. Moreover, it is worth noting that the problem size is much larger in their subtraction task than ours, and the operation effects may be confounded with the problem size effects. Given that both the operations and problem sizes are distinct from Keller & Menon (2009) and Pletzer (2016) as well as our studies, it is challenging to directly generalize the results.

Should operation and problem size confound with the sex/gender effect in brain response profiles, more extensive investigations are needed. Most studies did not directly compare the effect of problem size between sexes/genders on the brain activation profiles. Instead, problem size measured varies across studies with distinct indices and operations (Campbell & Xue, 2001; Grabner et al., 2007; Stanescu-Cosson et al., 2000). Our study intended to tackle this issue by systematically manipulating problem size, and provide a genuine effect of problem complexity in interpreting sex/gender differences in the brain response profiles.

4.4 Implications for using neuroimaging studies to understand sex/gender difference

The current findings highlight that the predictions obtained from behavioral performance may not always be appropriate to characterize brain configuration. This is illustrated by females and males engaging distinct brain response profiles even when their elicited behavioral performances remained the same. Insomuch of this assumption, it can be doubted that the previous observations of null results on sex/gender differences are likely underestimated. Behavioral assessments may not always secure such a level of cognitive processes. As a result, neuroimaging facilities, in contrast, have a strong potential to provide useful knowledge that is unseen in behavioral outcomes alone. Therefore, it is of crucial importance to provide unique perspectives using state-of-the-art neuroimaging techniques to understand biological sex differences in the human brain.

In view of our findings, it should be noted that males and females engage different response profiles of neural resources that can be influenced by problem solving strategies and affective factors to maintain parallel performance, indicating that differences in neural resource recruitment can be regulated by strategies and the consequence of other psychosocial factors (Taddei et al., 2022). These results suggested that the underrepresentation of females in math-related fields is more likely due to being blocked by psychological traits rather than inability. We propose that the strategy of supporting female students in their personal choice of math-related fields should likely focus on remediating these mental obstacles rather than providing prolonged instructions over the school mathematical materials. Should instructional practices be emphasizing elevating positive math attitude and more efficient strategies.

4.4 Conclusion

Over the past decades, cognitive and neural imaging studies have gained considerable insight into uncovering sex/gender differences in the mechanisms of learning. This work has led to advances in exploring the biological underpinnings of individual differences. However, direct manipulation of problem types during mathematical problem solving had not been systematically investigated. Our study emphasizes the importance of a linear task design in probing brain response profiles. Our findings revealed that, for the first time, problem complexity effects were markedly more prominent for females than males. On the other hand, using machine learning approach, we demonstrated that the fMRI signal profiles of the complexity are discriminative of the individual’s biological sex label. These results suggested that females and males take different but equivalently successful neural pathways to accomplish mathematical achievement. Further questions are raised, such as the effect of problem type, strategy selection, and the developmental progression. Future studies investigating potential neural mechanisms of when and how certain factors influence children's developing mathematical knowledge would improve the quality of school instruction and methods of teaching mathematics.

Reference

Ashcraft, M. H., & Krause, J. A. (2007). Working memory, math performance, and math anxiety. *Psychon Bull Rev, 14*(2), 243-248.

Behrens, T. E. J., Hunt, L. T., & Rushworth, M. F. S. (2009). The computation of social behavior. *Science, 324*(5931), 1160-1164. doi:10.1126/science.1169694

Behrens, T. E. J., Hunt, L. T., Woolrich, M. W., & Rushworth, M. F. S. (2008). Associative learning of social value. *Nature, 456*(7219), 245-249. doi:10.1038/nature07538

Burdina, G. M., Krapotkina, I. E., & Nasyrova, L. G. (2019). Distance learning in elementary school classrooms: An emerging framework for contemporary practice. *International Journal of Instruction, 12*(1), 1-16. doi:10.29333/iji.2019.1211a

Butterworth, B., Varma, S., & Laurillard, D. (2011). Dyscalculia: from brain to education. *Science, 332*(6033), 1049-1053. doi:10.1126/science.1201536

Dikker, S., Wan, L., Davidesco, I., Kaggen, L., Oostrik, M., McClintock, J., . . . Poeppel, D. (2017). Brain-to-Brain Synchrony Tracks Real-World Dynamic Group Interactions in the Classroom. *Curr Biol, 27*(9), 1375-1380. doi:10.1016/j.cub.2017.04.002

Ekstrom, R. B., French, J. W., & Harman, H. H. (1976). *Mual for kit of factor-referenced cognitive tests*. Princeton, NJ: Educational Testing Service.

Funane, T., Kiguchi, M., Atsumori, H., Sato, H., Kubota, K., & Koizumi, H. (2011). Synchronous activity of two people's prefrontal cortices during a cooperative task measured by simultaneous near-infrared spectroscopy. *Journal of Biomedical Optics, 16*(7), 077011.

Geary, D. C., Hoard, M. K., Nugent, L., & Bailey, D. H. (2013). Adolescents' functional numeracy is predicted by their school entry number system knowledge. *PLoS One, 8*(1), e54651. doi:10.1371/journal.pone.0054651

Haßler, B., Major, L., & Hennessy, S. (2016). Tablet use in schools: a critical review of the evidence for learning outcomes. *Journal of Computer Assisted Learning, 32*(2), 139-156. doi:10.1111/jcal.12123

Hari, R., Himberg, T., Nummenmaa, L., Hämäläinen, M., & Parkkonen, L. (2013). Synchrony of brains and bodies during implicit interpersonal interaction. *Trends in Cognitive Sciences, 17*(3), 105-106. doi:https://doi.org/10.1016/j.tics.2013.01.003

Hari, R., & Kujala, M. V. (2009). Brain Basis of Human Social Interaction: From Concepts to Brain Imaging. *Physiological Reviews, 89*(2), 453-479. doi:10.1152/physrev.00041.2007

Hu, Y., Hu, Y., Li, X., Pan, Y., & Cheng, X. (2017). Brain-to-brain synchronization across two persons predicts mutual prosociality. *Social Cognitive and Affective Neuroscience, 12*(12), 1835-1844. doi:10.1093/scan/nsx118

Koehne, S., Hatri, A., Cacioppo, J. T., & Dziobek, I. (2016). Perceived interpersonal synchrony increases empathy: Insights from autism spectrum disorder. *Cognition, 146*, 8-15. doi:10.1016/j.cognition.2015.09.007

Levy, J., Goldstein, A., & Feldman, R. (2017). Perception of social synchrony induces mother-child gamma coupling in the social brain. *Soc Cogn Affect Neurosci, 12*(7), 1036-1046.

doi:10.1093/scan/nsx032
Li, C., & Lalani, F. (2020). The COVID-19 pandemic has changed education forever.
Lindenberger, U., Li, S.-C., Gruber, W., & Müller, V. (2009). Brains swinging in concert: cortical phase synchronization while playing guitar. *BMC Neuroscience, 10*(1), 22. doi:10.1186/1471-2202-10-22
Liu, N., Mok, C., Witt, E. E., Pradhan, A. H., Chen, J. E., & Reiss, A. L. (2016). NIRS-Based Hyperscanning Reveals Inter-brain Neural Synchronization during Cooperative Jenga Game with Face-to-Face Communication. *Frontiers in Human Neuroscience, 10*(82).
Lyons, I. M., & Beilock, S. L. (2012). When math hurts: math anxiety predicts pain network activation in anticipation of doing math. *PLoS One, 7*(10), e48076. doi:10.1371/journal.pone.0048076
Maloney, E. A., & Beilock, S. L. (2012). Math anxiety: who has it, why it develops, and how to guard against it. *Trends Cogn Sci, 16*(8), 404-406. doi:10.1016/j.tics.2012.06.008
Mammarella, I. C., Hill, F., Devine, A., Caviola, S., & Szucs, D. (2015). Math anxiety and developmental dyscalculia: A study on working memory processes. *J Clin Exp Neuropsychol, 37*(8), 878-887. doi:10.1080/13803395.2015.1066759
Nozawa, T., Sakaki, K., Ikeda, S., Jeong, H., Yamazaki, S., Kawata, K. H. d. S., . . . Kawashima, R. (2019). Prior physical synchrony enhances rapport and inter-brain synchronization during subsequent educational communication. *Scientific Reports, 9*(1), 12747. doi:10.1038/s41598-019-49257-z
OECD. (2010). *The High Cost of Low Educational Performance: The Long-Run Economic Impact of Improving Educational Outcomes*.
OECD. (2013). *PISA 2012 Results (Volume III): Ready to learn: Students' engagement, drive and self-beliefs*.
Osaka, N., Minamoto, T., Yaoi, K., Azuma, M., Shimada, Y. M., & Osaka, M. (2015). How Two Brains Make One Synchronized Mind in the Inferior Frontal Cortex: fNIRS-Based Hyperscanning During Cooperative Singing. *Frontiers in Psychology, 6*(1811). doi:10.3389/fpsyg.2015.01811
Pan, Y., Novembre, G., Song, B., Li, X., & Hu, Y. (2018). Interpersonal synchronization of inferior frontal cortices tracks social interactive learning of a song. *NeuroImage, 183*, 280-290. doi:10.1016/j.neuroimage.2018.08.005
Prensky, M. (2007). Digital Game-Based Learning. New York: McGraw-Hill.
Ramirez, G., Gunderson, E. A., Levine, S. C., & Beilock, S. L. (2013). Math Anxiety, Working Memory, and Math Achievement in Early Elementary School. *Journal of Cognition and Development, 14*(2), 187-202. doi:10.1080/15248372.2012.664593
Richland, L. E., Zur, O., & Holyoak, K. J. (2007). Mathematics. Cognitive supports for analogies in the mathematics classroom. *Science, 316*(5828), 1128-1129. doi:10.1126/science.1142103
Sher, A. (2009). Assessing the relationship of student-instructor and student-student interaction

to student learning and satisfaction in web-based online learning environment. *Journal of Interactive Online Learning*.

Sänger, J., Müller, V., & Lindenberger, U. (2012). Intra- and interbrain synchronization and network properties when playing guitar in duets. *Frontiers in Human Neuroscience, 6*(312). doi:10.3389/fnhum.2012.00312

Strang, K. (2013). Cooperative learning in graduate student projects: Comparing synchronous versus asynchronous collaboration. *Journal of Interactive Learning Research, 24*(4), 447-464.

Suarez-Pellicioni, M., Nunez-Pena, M. I., & Colome, A. (2015). Math anxiety: A review of its cognitive consequences, psychophysiological correlates, and brain bases. *Cogn Affect Behav Neurosci*. doi:10.3758/s13415-015-0370-7

Supekar, K., Iuculano, T., Chen, L., & Menon, V. (2015). Remediation of Childhood Math Anxiety and Associated Neural Circuits through Cognitive Tutoring. *J Neurosci, 35*(36), 12574-12583. doi:10.1523/JNEUROSCI.0786-15.2015

Valdesolo, P., & Desteno, D. (2011). Synchrony and the social tuning of compassion. *Emotion, 11*(2), 262-266. doi:10.1037/a0021302

Watts, L. (2016). Synchronous and asynchronous communication in distance learning: a review of the literature. *Quarterly Review of Distance Education, 17,* 23

Wechsler, D. (2004). *The Wechsler intelligence scale for children - fourth edition*. London: Pearson Assessment.

Wiltermuth, S. S., & Heath, C. (2009). Synchrony and Cooperation. *Psychological Science, 20*(1), 1-5. doi:10.1111/j.1467-9280.2008.02253.x

Yang, D. Y., Rosenblau, G., Keifer, C., & Pelphrey, K. A. (2015). An integrative neural model of social perception, action observation, and theory of mind. *Neurosci Biobehav Rev, 51*, 263-275. doi:10.1016/j.neubiorev.2015.01.020

林宜慧. (2007). *台南地區國一學生參加校外數學補習對其學習的影響*. 國立高雄師範大學, 高雄市.

洪儷瑜, & 連文宏. (2015). *基本數學核心能力測驗*. 台北: 中國行為科學社.

徐政業. (2008). *學校與補習班數學教育情況及其對學生數學學習影響之比較*. (碩士). 國立臺北師範大學, 台北市.

張雅俐. (2011). *國小學童數學學業成就、數學補習經驗與數學學習策略關係之研究*. (碩士). 國立臺南大學, 台南市.

許綺婷. (2001). *探討國三學生對補習班與學校教學的看法及其與基本學測數學科之表現的關係*. (碩士). 國立台灣師範大學, 台北市.

科技部補助專題研究計畫出席國際學術會議心得報告

日期：112 年 1 月 8 日

<table>
<tr><td>計畫編號</td><td colspan="3">MOST110－2629－H－004－001</td></tr>
<tr><td>計畫名稱</td><td colspan="3">以磁振造影研究探討數學能力之性別差異</td></tr>
<tr><td>出國人員姓名</td><td>陳心喻</td><td>服務機構
及職稱</td><td>國立政治大學心理學系
碩士班研究生</td></tr>
<tr><td>會議時間</td><td>111 年 4 月 23 日
至
111 年 4 月 26 日</td><td>會議地點</td><td>美國舊金山</td></tr>
<tr><td>會議名稱</td><td colspan="3">(英文) Cognitive Neuroscience Society Annual Meeting 2022</td></tr>
<tr><td>發表題目</td><td colspan="3">(英文) Age-related differences of inhibitory control engagement underlying arithmetic abilities between children and adolescent</td></tr>
</table>

一、參加會議經過

由於疫情嚴峻，本年度會議採實體與線上並行的方式進行，而學生在四月二十三日到二十六日間以線上形式參與海報展示 (poster sessions)，在會議前，主辦方提供網站連結供學生上傳海報電子檔及報告影片，在會議期間透過線上參與的與會者皆可透過該網站觀看上述資料。在會議期間，學生除了可以在該網站瀏覽眾多投稿者的海報外，還能收看演講及工作坊，其中學生對一位投稿者海報中提到的技術十分感興趣，故有透過該投稿者提供的電郵地址與其取得聯繫並進行延伸的討論。

二、與會心得

本次雖然無法實體與會，但藉由線上參與的方式，還是可以觀看到其他投稿者的海報發表，除了能夠拓增在研究上的眼界外，同時也能學習到海報的製作與編排技巧，讓報告人大有獲益。在眾多海報中，與學生研究較為相關的是用 fMRI 技術探討認知的發表，但由於認知的範疇之多之廣，在會議中能夠看到與自身研究領域不太相同的主題，如 social learning、brain injury and diseases 和 emotion 等等，對學生來說皆十分有趣，也體會到雖然這些主題看似與學生的研究不同，但絕非完全沒有相關，相信對日後研究上可以提供另一個角度的思考，著實助益良多。另外學生又特別感興趣的是關注於 white matter tractography 的研究，因此透過其他報告者的海報發表，學生得以觀摩他人的研究方法，並以通信的方式向原作者提出問題來進行交流與討論，對學生來說也訓練了英文學術通信的技巧，並領悟到要與國際接軌，才能激發更多新穎且專業的研究想法，因此十分感謝國科會補助經費還有國立政治大學的行政協助，讓研究生能得到這樣的經驗，對未來生涯發展也有相當的鼓舞與激勵作用。

三、發表論文全文或摘要

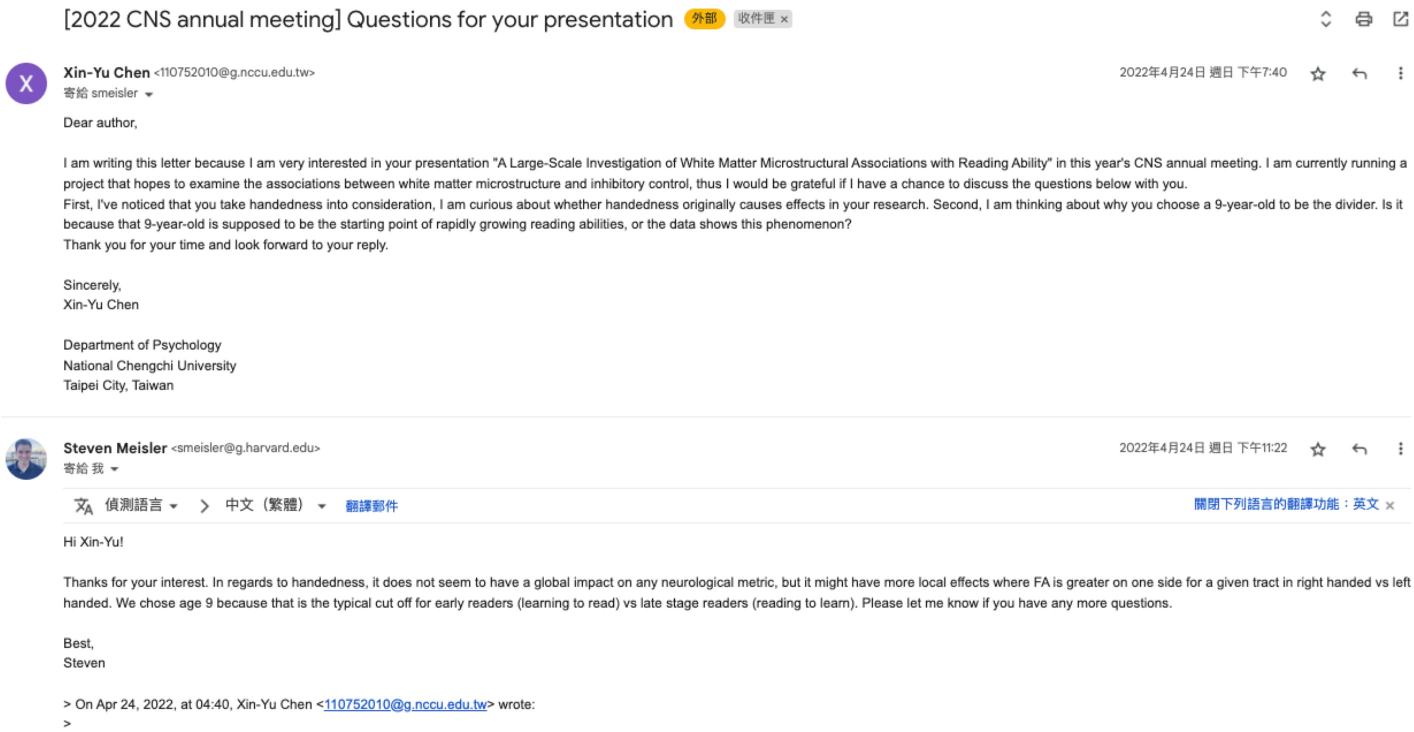
於報告書最後附上。

四、建議

對研究生而言，參與國際會議是難得且珍貴的機會，不但能與國外的研究者進行經驗交流，也能藉由與其他研究者的討論更釐清自己的研究，如有機會應把握並多多參與，以提升自己的視野，是非常好的學習。

五、其他

圖一：線上參與會議網站



[2022 CNS annual meeting] Questions for your presentation 外部 收件匣 ×

Xin-Yu Chen <110752010@g.nccu.edu.tw> 2022年4月24日 週日 下午7:40
寄給 smeisler

Dear author,

I am writing this letter because I am very interested in your presentation "A Large-Scale Investigation of White Matter Microstructural Associations with Reading Ability" in this year's CNS annual meeting. I am currently running a project that hopes to examine the associations between white matter microstructure and inhibitory control, thus I would be grateful if I have a chance to discuss the questions below with you.
First, I've noticed that you take handedness into consideration, I am curious about whether handedness originally causes effects in your research. Second, I am thinking about why you choose a 9-year-old to be the divider. Is it because that 9-year-old is supposed to be the starting point of rapidly growing reading abilities, or the data shows this phenomenon?
Thank you for your time and look forward to your reply.

Sincerely,
Xin-Yu Chen

Department of Psychology
National Chengchi University
Taipei City, Taiwan

Steven Meisler <smeisler@g.harvard.edu> 2022年4月24日 週日 下午11:22
寄給 我

偵測語言 > 中文（繁體） 翻譯郵件 關閉下列語言的翻譯功能：英文 ×

Hi Xin-Yu!

Thanks for your interest. In regards to handedness, it does not seem to have a global impact on any neurological metric, but it might have more local effects where FA is greater on one side for a given tract in right handed vs left handed. We chose age 9 because that is the typical cut off for early readers (learning to read) vs late stage readers (reading to learn). Please let me know if you have any more questions.

Best,
Steven

> On Apr 24, 2022, at 04:40, Xin-Yu Chen <110752010@g.nccu.edu.tw> wrote:
>

圖二：與海報原作者討論之往返信件

Abstract

Inhibitory control (IC), the capacity to suppress an inappropriate prepotent response, plays a crucial role in building foundational cognitive skills, especially during the early stage of development. Although neuroimaging studies have provided abundant evidence that brain responses associated with inhibitory control are consistently implicated in children's mathematical learning, how IC develops across school stage into adolescence is still poorly understood. In this study we investigate this issue using fMRI methods. Brain responses of fifty-two children (ages 7-13) and twenty-two adolescents (ages 13-18) were acquired while they performed an arithmetic task comprised by large and small problems in the MRI scanner. All participants were categorized as higher and lower IC groups by median split using the performance of a flanker task administered outside the MRI scanner. Voxel-wise three-way ANOVA with problem size (large, small) as a with-subject factor and age (children, adolescents) as well as IC (high, low) as between-subject factors were examined across the whole brain. The results revealed three-way interaction, with children with higher IC show stronger activations in the frontal-parietal regions, including middle frontal gyrus and intraparietal sulcus, compared to those with lower IC. In contrast, adolescents with higher IC show more deactivations in default mode network, including precuneus, angular gyrus, and ventromedial prefrontal cortex, than the lower IC group. These results suggested that the cognitive and neural mechanisms of inhibitory control underlying arithmetic learning develops across essential school stages. Our study therefore provides insights into uncovering the biological underpinnings of the maturation of cognitive skill acquisition.

EXECUTIVE PROCESSES: Monitoring & inhibitory control

Age-related differences of inhibitory control engagement underlying arithmetic abilities between children and adolescents

NATIONAL CHENGCHI UNIVERSITY

Xin-Yu CHEN[1], Chan-Tat NG[1], Ting-Ting CHANG[1,2]

[1]Department of Psychology, National Chengchi University, Taipei City, Taiwan; [2]Research Center for Mind, Brain, and Learning, National Chengchi University, Taipei City, Taiwan.

大腦與學習實驗室 Brain and Learning Lab

Background

- Previous studies have showed that brain responses associated with inhibitory control (IC) are consistently implicated in children's math learning (Sarama, & Germeroth, 2016) .
- IC is particularly crucial for learning and school attainments.
- How IC develops across elementary school into high school stage is still poorly understood.
- Here we investigate this issue by comparing elementary school and high school students using fMRI.

Methods & Results

- Participants: 52 children from grade 1 to grade 6 (ages 7-13) and 22 adolescents from junior and senior high school (ages 13-18).

A. Flanker task (out-of-scanner)

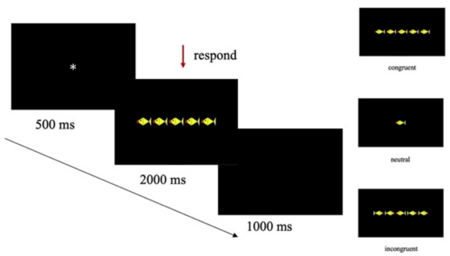


Figure 1. Procedure of Flanker task. Participants were asked to identify the direction of the central target.

Methods & Results

B. Arithmetic task (in-scanner)

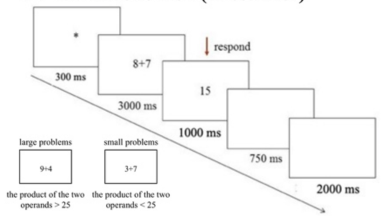


Figure 2. Procedure of Arithmetic task. Participants calculated the presented problem and verified whether the answer corresponded to the following number.

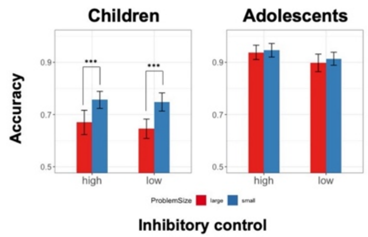


Figure 3. Accuracy of the in-scanner arithmetic task. ***$p < .001$.

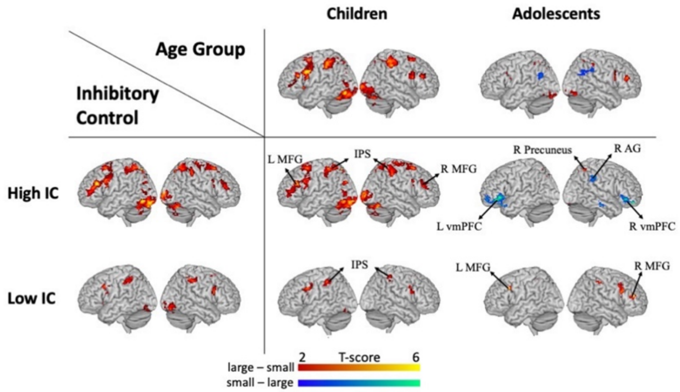


Figure 4. Brain regions that showed problem size effect in children/adolescents and high/low inhibitory control (IC). All participants were categorized as higher and lower IC groups by median split using the performance of a flanker task.

Conclusion

- Problem size effect showed in both high and low IC participants in children's group, but not in adolescents' group.
- Children with higher IC showed stronger problem size effect in frontal-parietal regions including MFG and IPS as well as ventrotemporal occipital cortex than those with lower IC.
- Adolescents with higher IC showed more deactivations in the DMN, which consisted of precuneus, AG, and vmPFC.
- Our study suggested that the cognitive and neural mechanisms of inhibitory control underlying arithmetic learning developed across essential school stages.

References

Clements, D. H., Sarama, J., & Germeroth, C. (2016). Learning executive function and early mathematics: Directions of causal relations. *Early Childhood Research Quarterly, 36*, 79-90.

Acknowledgments

The research was supported by the grants from the Ministry of Science and Technology (MOST108-2410-H-004-101-MY2; MOST106-2410-H-004-066-MY3; MOST105-2511-S-004-001-MY3)

Contact information

Xin-Yu CHEN 110752010@g.nccu.edu.tw

110年度專題研究計畫成果彙整表

<table>
<tr><td colspan="3">計畫主持人：張葶葶</td><td colspan="3">計畫編號：110-2629-H-004-001-</td></tr>
<tr><td colspan="6">計畫名稱：以磁振造影研究探討數學能力之性別差異</td></tr>
<tr><td colspan="3">成果項目</td><td>量化</td><td>單位</td><td>質化
（說明：各成果項目請附佐證資料或細項說明，如期刊名稱、年份、卷期、起訖頁數、證號...等）</td></tr>
<tr><td rowspan="6">國內</td><td rowspan="6">學術性論文</td><td>期刊論文</td><td>0</td><td rowspan="2">篇</td><td></td></tr>
<tr><td>研討會論文</td><td>0</td><td></td></tr>
<tr><td>專書</td><td>0</td><td>本</td><td></td></tr>
<tr><td>專書論文</td><td>0</td><td>章</td><td></td></tr>
<tr><td>技術報告</td><td>0</td><td>篇</td><td></td></tr>
<tr><td>其他</td><td>0</td><td>篇</td><td></td></tr>
<tr><td rowspan="6">國外</td><td rowspan="6">學術性論文</td><td>期刊論文</td><td>1</td><td rowspan="2">篇</td><td>Ting-Ting Chang*; Nai-Feng Chen; Yang-Teng Fan, 2022.09, 'Uncovering sex/gender differences of arithmetic in the human brain: insights from fMRI studies,' Brain and Behavior, Vol.12, No.10, pp.e2775.(SCI)</td></tr>
<tr><td>研討會論文</td><td>0</td><td></td></tr>
<tr><td>專書</td><td>0</td><td>本</td><td></td></tr>
<tr><td>專書論文</td><td>0</td><td>章</td><td></td></tr>
<tr><td>技術報告</td><td>0</td><td>篇</td><td></td></tr>
<tr><td>其他</td><td>0</td><td>篇</td><td></td></tr>
<tr><td rowspan="10">參與計畫人力</td><td rowspan="5">本國籍</td><td>大專生</td><td>0</td><td rowspan="10">人次</td><td></td></tr>
<tr><td>碩士生</td><td>2</td><td>陳心喻、黃如君</td></tr>
<tr><td>博士生</td><td>0</td><td></td></tr>
<tr><td>博士級研究人員</td><td>1</td><td>陳乃鳳</td></tr>
<tr><td>專任人員</td><td>1</td><td>范惠君</td></tr>
<tr><td rowspan="5">非本國籍</td><td>大專生</td><td>3</td><td>朱莫亞、程妍慈、劉奕姍</td></tr>
<tr><td>碩士生</td><td>0</td><td></td></tr>
<tr><td>博士生</td><td>1</td><td>伍贊達</td></tr>
<tr><td>博士級研究人員</td><td>0</td><td></td></tr>
<tr><td>專任人員</td><td>0</td><td></td></tr>
<tr><td colspan="3">其他成果
（無法以量化表達之成果如辦理學術活動、獲得獎項、重要國際合作、研究成果國際影響力及其他協助產業技術發展之具體效益事項等，請以文字敘述填列。）</td><td colspan="3">本計畫除已於Brain and Behavior期刊發表"Uncovering sex/gender differences of arithmetic in the human brain: insights from fMRI studies"一文之外，另已投稿"Problem complexity modulates brain responses of mental arithmetic in females but not in males"一文於期刊Journal of Neuroscience Research，目前已修改一次，正在第二次審查中。發表文</td></tr>
</table>

	章研究成果請見報告內文。