

ORIGINAL ARTICLE

The Neural Correlates of Computational Thinking: Collaboration of Distinct Cognitive Components Revealed by fMRI

Shan Xu¹, Yan Li¹ and Jia Liu²¹Faculty of Psychology, Beijing Normal University, Beijing 100875, China and ²Department of Psychology, Tsinghua Laboratory of Brain and Intelligence, Tsinghua University, Beijing 100086, China

Address correspondence to Shan Xu, Faculty of Psychology, Beijing Normal University, Beijing 100875, China. Email: shan.xu@bnu.edu.cn or Jia Liu, Department of Psychology, Tsinghua Laboratory of Brain and Intelligence, Tsinghua University, Beijing 100086, China. Email: liujiaTHU@tsinghua.edu.cn

Abstract

Recent technical advance attracts great attention to the promotion of programming skills, in particular, and computational thinking (CT), in general, as a new intellectual competency. However, the understanding of its cognitive substrates is limited. The present study used functional magnetic resonance imaging to examine the neural correlates of programming to understand the cognitive substrates of CT. Specifically, magnetic resonance imaging signals were collected while the participants were mentally solving programming problems, and we found that CT recruited distributed cortical regions, including the posterior parietal cortex, the medial frontal cortex, and the left lateral frontal cortex. These regions showed extensive univariate and multivariate resemblance with arithmetic, reasoning, and spatial cognition tasks. Based on the resemblance, clustering analyses revealed that cortical regions involved in CT can be divided into Reasoning, Calculation, Visuospatial, and Shared components. Further, connectivity increased during programming within the CT network constructed by these four components and decreased between the CT network and other cortical regions. In sum, our study revealed the cognitive components underlying CT and their neural correlates and further suggests that CT is not a simple sum of parallel cognitive processes, but a composite cognitive process integrating a set of intellectual abilities, particularly those in the science, technology, engineering, and math domains.

Key words: cognitive components, computational thinking, fMRI, programming

Introduction

In recent years, computer science plays an increasingly critical and ubiquitous role in human life. It brings one of the most significant challenges to human intelligence in our time, that is, to adopt to and make the best use of these new advances of computer science and the vast amount of information generated by the newly acquired computational power. The ability to do so has been deemed of great value nowadays, and the efficient training and evaluating becomes a great need of society. Specifically, the growing educational need underscores the urgency in characterizing the cognitive and neural mechanisms

behind the acquisition and use of computational techniques and computer programming skills. Initially, the concept of computational thinking (CT) is proposed as mental skills and thinking habits that people develop through their work in computing disciplines, such as programming (Papert 1990; Tedre and Denning 2016), or a synthesis of cognitive abilities in problem-solving processes that drive programming skills (Denning 2017). A less concentrated view of CT weakens the link between CT and programming tasks and disciplines, but defines CT as a cross-disciplinary literacy that is widely required in science, technology, engineering, and math (STEM) disciplines and daily life (Mannila et al. 2014; Weintrop et al. 2016). For instance, CT

is surmised as the ability to “think like a computer scientist” (Wing 2006), referring to solving problems, designing systems, and understanding human behavior by drawing on concepts that are fundamental in computer science (Wing 2006). It is considered to include (but not limited to): formulating problems for computational solutions, logically organizing and analyzing data, abstractions including models and simulations, algorithmic thinking, evaluation for efficiency and correctness, and generalizing and transferring to other domains (CSTA and ISTE 2011). Admittedly, in contrast to the increased attention that CT receives in the field of education and cognitive sciences, there still lacks an empirically proven consensus on a formal generic definition of CT (Grover et al. 2014; Kalelioglu et al. 2016), and the cognitive substrates of CT are also not clear. As many other core competencies, such as reading, writing and arithmetic, CT is a product of an adaptation to human civilization. Therefore, it is unlikely that a new cognitive module would emerge for this newly arising challenge; rather, CT presumably arises from the recruitment, adaptation, or even integration of phylogenetically older cognitive abilities. Indeed, several cognitive components of CT, such as abstraction and automation, decomposition, algorithmic thinking, and representing data as models and simulations have been proposed based on theoretical consideration (e.g., Wing 2006, 2008; CSTA and ISTE 2011); however, these intuitions lack empirical evidence, and it is unclear how these cognitive components are integrated as a cohesive unity. Therefore, the application of CT is largely limited in practice.

A most representative task for CT is programming, which has started to be included in school curriculum in countries across Europe, Asia, and North and South America (Fedorenko et al. 2019), and efforts have been made to understand the neural and cognitive correlates of programming. For instance, behavioral and neuroimaging correlation analyses have revealed that the behavioral performance of code review is correlated with the capacity and neural correlates of working memory (Crk et al. 2016; Yeh et al. 2017; Baum et al. 2019). Further, functional magnetic resonance imaging (fMRI) studies on code review and comprehension have identified the involvement of distributed cortical regions, including the inferior and the middle frontal gyri, the middle temporal gyrus, the lateral occipital regions, and the inferior parietal lobule, which are typically associated with language processing, high-level visual processing and reasoning (Prabhakaran et al. 1997; Fedorenko and Thompson-Schill 2014; Hobeika et al. 2016) as well as the insula and the medial frontal cortex (Duraes et al. 2016; Castelhana et al. 2019), which are typically associated with error monitoring (Sharp et al. 2010; Bastin et al. 2016). However, programming is a complex cognitive task (Dalbey and Linn 1985), and yet most existing studies focus on the comprehension stage of programming, such as code comprehension or code review, which reflects the mastery or proficiency of particular coding language or syntax, but may not fully address one of the main functions of programming, that is, using it as a tool for problem-solving. The present study, therefore, intended to investigate the neural correlates of programming using a task relying more on its problem-solving and algorithm generation aspects.

Specifically, the present study adopted a set of simple programming problems from programming training, and participants equipped with basic programming skills mentally solved these problems in the scanner (Fig. 1). Besides, four benchmark tasks on cognitive processes that may underlie CT, that is, language, arithmetic, reasoning, and spatial cognition, chosen based on previous theoretical proposals, were also

included (Fig. 2). The language ability was examined because of its hypothetic relation to programming (Fedorenko et al. 2019) as suggested by the association between the language abilities and programming practice or the master of programming language (Pepler and Warschauer 2011; Kazakoff and Bers 2014) and the activation of language-related cortical regions during code comprehension (Siegmund et al. 2014). The arithmetic task was chosen as one representational aspect of the STEM disciplines that computer programming is presumably associated with (Fedorenko et al. 2019). The spatial cognition task was chosen as another representative aspect of the STEM disciplines, and it has been suggested related to CT (Roman-Gonzalez et al. 2017). Finally, programming has frequently been discussed in the framework of reasoning and problem-solving theoretically (Dalbey and Linn 1985; Fedorenko et al. 2019), which is confirmed by a recent psychometrical study (Roman-Gonzalez et al. 2017). With the neural correlates of these four tasks as benchmarks, we investigated the cognitive substrates of CT and its potential underlying cognitive components with both univariate activation and multi-variate pattern analyses. We exploratorily identified functional modules from the neural correlates of programming with a data-driven clustering analysis. We reasoned that if such functionally dissociable modules had existed, their activation profiles and pattern similarity profiles would have varied across the experimental tasks. Then, with functional connectivity analysis among these underlying modules, we examined how they are integrated to form a cohesive unity of CT.

Besides the domain-specific functions, some domain-general cognitive functions, such as working memory, executive control, and decision-making, may also be recruited for CT. For instance, CT performance has been found to correlate with working memory (Ambrósio et al. 2014), and the above-mentioned studies on code review have also suggested the involvement of working memory and executive control in that representative CT task. In addition, decision-making has been proposed relevant to CT in line with the theoretical proposal of evaluation, one of the core components of CT that refers to the finding of the best solution for making decisions about good use of resources (CSTA and ISTE 2011; Selby and Woollard 2013). Nevertheless, because the domain-general functions are recruited in many tasks unrelated to CT, here, we focused on the domain-specific functions that are closely related to CT.

Materials and Methods

Participants

Twenty participants (right-handed neurologically normal volunteers with normal or corrected-to-normal vision) completed the experiment; another two were initially recruited but failed to complete the experiment and therefore were not included in the analysis. The participants were screened before recruitment to ensure they either had taken programming-related courses or had experience in programming. The targeted sample size (20 valid participants) was decided a priori to be similar to previous studies using similar benchmark tasks (e.g., Fedorenko et al. 2012, $n = 13$ – 16 ; Zhou et al. 2018, $n = 24$; Amalric and Dehaene 2016, $n = 15$), and participant recruitment stopped immediately after the targeted sample size was met. The study was approved by the Institutional Review Board of Beijing Normal University (BNU). Written informed consent was obtained

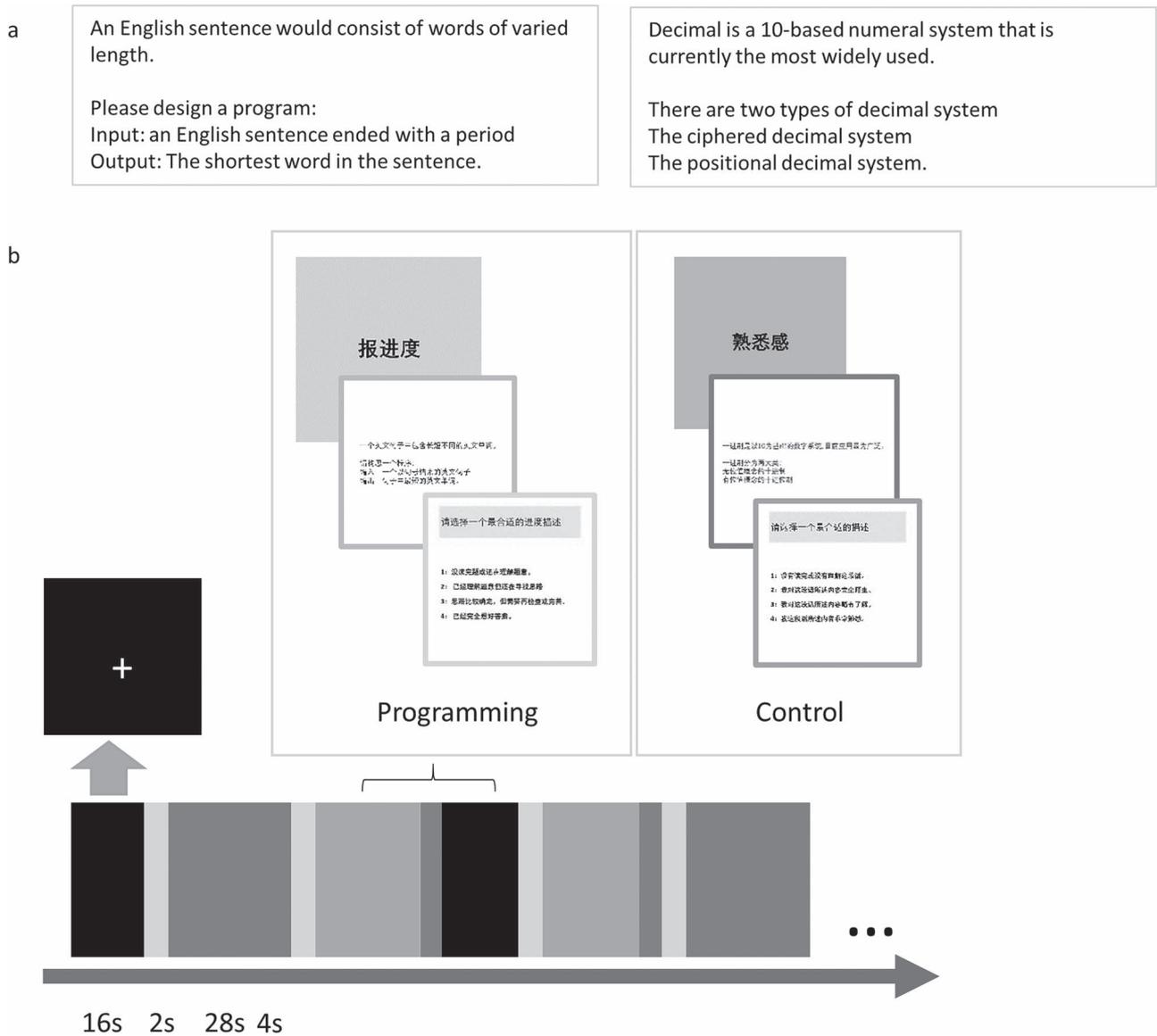


Figure 1. The illustration of the programming task. (a) The English translation of two examples of the programming items (left) and the control items (right). (b) The procedure of the scanning session (bottom) and the examples of the programming condition (upper left) and its control condition (upper right). From top to bottom, the three screens in each gray frame are the task probe, the programming item, and the response prompt, respectively. Both types of trials started with a task probe for 2 s, followed by the programming item for 28 s and the response prompt for 4 s.

from all participants before they took part in the experiment, and the participants received money for their time.

Procedure

The trial sequences of the experimental tasks were generated and presented using Matlab R2016b (The MathWorks Inc.) with Psychtoolbox 3. Each participant participated in five tasks, consisting of one core task, the programming task, and four benchmark tasks (see below for details) in the scanner, and filled a short questionnaire about their educational and practical experience in programming after they completed the programming task. The tasks were divided into three scanning sessions conducted on three separate days. The order of the five tasks was

not counterbalanced across subjects. For each task, the participants received instruction, got familiar with the requirement of each condition, and completed a practice run before entering the scanner.

Programming

The materials used in the programming condition of the programming task consisted of 24 programming problems adopted from elementary practice problems of International Olympiad Informatics and other advanced programming competitions for secondary school students. We collected the original problems from the Internet, translated those originally presented in English into Chinese, removed those containing graphs, figures and chunky codes or long paragraphs, and shortened and

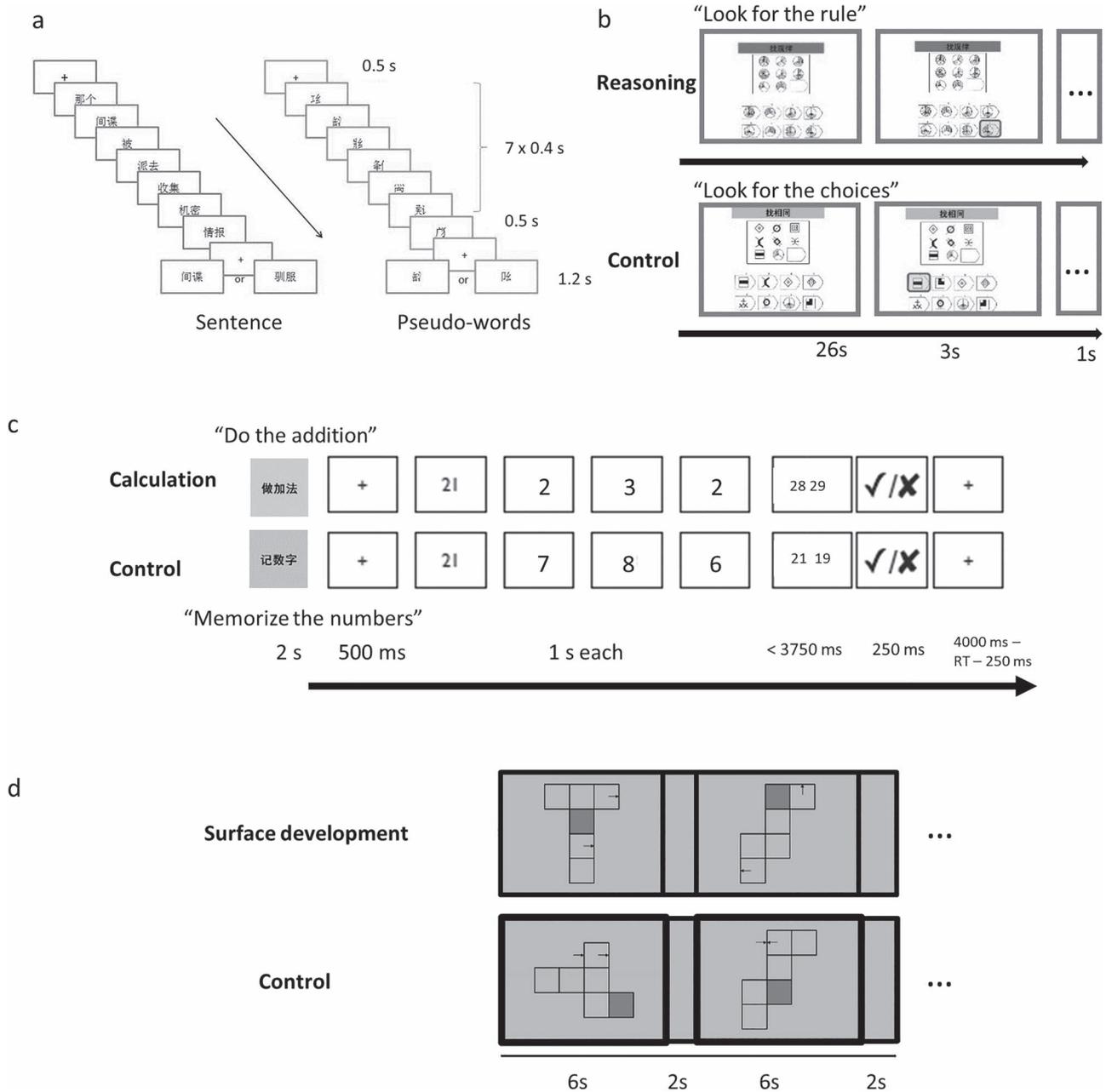


Figure 2. The illustration of trials of the four benchmark tasks: language (a), reasoning (b), arithmetic (c), and spatial cognition (d) tasks. The translation of the task probes of the reasoning and arithmetic tasks were denoted in quotation marks.

reformatted each of the remaining problems into a one-to-three-line introductory statement, a programming probe (“Please design a program with:”), a one-line statement specifying the potential input, and a one-line statement describing the desired output. For the control condition, we generated knowledge entries with the length, the format, and the use of symbols matched with the materials of the programming condition. Each entry started with a one-to-three-line introductory statement, followed by a brief conjunctive clause (e.g., “in other words”), and two one-to-two-line deliberations of the concept introduced in the introductory statement. The items were presented at the center of a light gray (204, 204, 204 RGB) background subtended

18.7° × 18.7° of visual angle. Other stimuli included a fixation cross subtended 0.6° × 0.6° of visual angle, two task probes subtended 18.7° × 18.7° of visual angle, and two responses prompt subtended 18.7° × 18.7° of visual angle (see Fig. 1 for an example item of each condition and their English translation).

The programming task was conducted in the scanner, divided into six runs, each with four trials in the programming condition and four in the control condition. Pairs of trials of the same condition interleaved with trials in the other condition in an ABBA manner. The trial order in a given run was identical across subjects, but the run order was balanced across participants. For half of the participants, the odd number runs started with a pair

of programming condition trials and the rest runs control trials, while the other half went through the six runs in the reversed order. The programming trials started with a programming task probe, 2 s presentation of “report progress” on a light green (197, 224, 181 RGB) background. In these trials, the participants were to read and consider the programming problem in the following 28 s and wait for the response prompt (“Please select the most appropriate statement”) and choose from a four-point scale to report their progress in programming. The scale was presented under the prompt, ranging from “1: Have not finished reading or do not understand” to “4: Ready with a solution.” The participants were required to press a corresponding key of the response boxes held in their hands to indicate their choice. They were instructed before the scanning that for programming they should consider at the algorithm level instead of specific code, and the solution of the problem means finding the algorithm instead of finishing the coding. The response prompt and the scale would remain presented for 4 s. The control trials started with the control task probe which read “report familiarity” on a light blue (189, 215, 238 RGB) background. In these trials, the participants were to read and consider whether they are familiar with the content of the knowledge entry in the following 28 s and wait for the response prompt and choose from a four-point scale ranging from “1: I have not finished reading or do not understand” to “4: I am familiar with all the information in the statement” (see Fig. 1 for the trial procedure as well as example materials). At the beginning and the end of each run, as well as between the fourth and the fifth trials of each run, 16-s null events were inserted during which only the fixation cross was presented. After the scan, we invited the participants to a debriefing session in which they were shown with a set of the tested items again and were required to immediately report their thoughts in solving this problem and the solution they had come up with during the scan. All the participants passed this debriefing by reporting reasonable solutions, or signs of heading toward them, without further consideration.

The Language Task

The language task was adopted from Fedorenko et al. (2010), but with materials (i.e., sentences and pseudo-characters) in Mandarin Chinese. Each participant completed four block-design functional runs. Each run contained three blocks of sentences and three blocks of strings of Chinese pseudo-characters, which were visually presented and separated by seven fixation blocks. For three participants, each fixation block lasted for 12 s and, for the rest, 16 s. The pseudo-characters were constructed of components of real characters, but they had no meaning or phonology of their own. Each block lasted for 30 s and consisted of six trials. Each trial contained one sentence that was formed by a string of seven Chinese real words or seven pseudo-characters, which were followed by a target word or pseudo-character. Participants were instructed to decide whether the target word or pseudo-character appeared in the preceding sentence or string of pseudo-characters and responded by pressing certain buttons (for more details on the paradigm, see Xu et al. 2015). Trials with pseudo-word stimuli were the control condition of this task.

The Arithmetic Task

The arithmetic task was adopted from Fedorenko et al. (2010) with an additional control condition. Each participant completed four blocked-design functional runs. In the arithmetic trials, the participants first see a task probe (“Do the addition”

on light-green background) presented for 1.5 s, then after a 0.5-s ISI with the fixation cross on the screen, they see one number (11–30) followed by three sequentially presented addends (of sizes 2–4 in half of the trials and 6–8 in the rest. The two kinds of trials were not differentiated in the analysis.) on light-gray background. Then the participants had to choose the correct sum in a two-choice, forced-choice question. After a response was made, brief feedback was shown on the screen to tell the participants whether they answered correctly. In the control trials, after the task probe (“Memorize the numbers” on light-blue background) and the same sequential presentation of four numbers, the participants were to select, from two alternatives, the one among the four preceding numbers. The trials were grouped into four-trial blocks. Each run consisted of six 40-s calculation blocks and six 40-s control blocks. For five participants, two 16-s null events were added in the beginning and after the ninth block of each run, for the rest of the participants, three 16-s null events were added at the beginning, the end, and after the sixth trial of the run. The order of the trials within each run was pseudo-randomized to balance the frequency of the numbers and correct choices. This order was kept identical across the participants.

The Raven Reasoning Task

Each participant completed five functional runs. The task consisted of a reasoning condition adopted from Raven’s APM and a control condition. In both conditions, the participants first saw a problem with a complex main figure, with a blank space in its right bottom corner. Below it were eight choices. The problems of the reasoning trials were chosen from Raven’s APM (problems 3–12 of set 1 and problems 23, 24, 27–30, 32–34, and 36 of set 2). The problems of the control condition were composed by rearranging the choices of other Raven’s APM items in the main figure and the choices in each problem. Above each problem, a task probe would be presented to indicate whether it is a reasoning trial (“Look for the rule” on a green background) or a control trial (“Look for the choices” on a fuchsia background). The participants were to consider which choice complies with the rule of figure arrangement and completes the main pattern (the reasoning condition), or to consider which choice was presented in the main pattern (the control condition). After 26 s, one of the choices would be highlighted by a red frame, and the participants were instructed to report whether this choice was the correct one within 3 s, then the next trial started after a 1-s ISI. Each run consisted of eight trials. Three 16-s null events were added in the beginning, the end, and after the fourth trial of the run. The order of the trials within each run was randomized and kept identical across participants. The order of conditions was similarly balanced within and across the runs as in the programming task. The order of runs was balanced between participants in the same manner as in the programming task.

The Surface Development Task

The surface development (SD) task was adopted from the mental folding task of Milivojevic et al. (2003). The task consisted of a SD condition and a control condition. In each trial, the participants were shown a black outline of six squares joined together, representing the faces of an unfolded cube. Two small arrows pointed to two of the sides of a square. In half of the trials, that is, the “match” trials, the pointed sides would meet if the squares were folded up into a cube. Trials in the SD and the control conditions differed in the total number of squares carried

along for each fold in order for the match–mismatch decision to be made. The control trials presented 1-square-carried stimuli since judgment of such stimuli can be made without mental folding (see Milivojevic et al. 2003 for more details). Each trial would present the outline stimuli for 6 s, followed by a 2 s blank screen, and the participants were instructed to indicate their judgment by pressing one of two buttons of the response box. The order of the trials within each run was pseudo-randomized and was kept identical across participants. Every four trials of the same condition would be grouped into a block, and the order of blocks was balanced in each run in the ABBA manner. For three participants, a 16-s null event was inserted before and after every block, while for the rest, the null events were inserted before and after every two blocks. The order of conditions within runs and the order of runs were balanced in the same manner as in the programming task. Each participant completed five or six block-design functional runs.

fMRI Data Acquisition

fMRI data were acquired using a Siemens 3T scanner (MAGEN-TOM Trio, a Tim system) with a 12-channel phased-array head coil at the Beijing Normal University Imaging Center for Brain Research, Beijing, China. Task-state fMRI (ts-fMRI) was acquired using a T2*-weighted echo-planar-imaging (EPI) sequence with a whole-brain protocol (TR = 2000 ms, TE = 30 ms, flip angle = 90°, and in-plane resolution = 3.1 × 3.1 × 3.5 mm, 33 contiguous interleaved slices). In addition, high-resolution T1-weighted images were acquired with a magnetization-prepared GRE sequence (MPRAGE: TR/TE/TI = 2530/3.39/1100 ms, flip angle = 7°, matrix = 256 × 256) for spatial registration. Earplugs were used to attenuate the scanner noise, and a foam pillow and extendable padded head clamps were used to restrain head motion. We also collected resting-state fMRI data, which were not analyzed in the present study. All the stimuli were projected onto a screen at the back of the scanner and were viewed from a distance of approximately 110 cm via a mirror placed on the head coil.

fMRI Data Analysis

Data preprocessing was performed with DPABI (Yan et al. 2016, <http://rfmri.org/dpabi>). The main preprocessing procedure was as follows: 1) transformation of DICOM files into NIFTI images, 2) slice timing, 3) head motion correction, 4) co-registration of the high-resolution T1-weighted structural images to the functional images, 5) segmentation, 6) spatial normalization to standard Montreal Neurological Institute (MNI) space, and resampling to 3 × 3 × 3 mm isotropic voxels, and 7) smoothing with a 4-mm full-width-half-maximum Gaussian kernel for univariate activation analysis. For multiple regression of activation pattern, we did not include smoothing in the preprocessing pipeline. We excluded runs with excessive head movement (>2 mm in any direction across a run) from further analysis.

The data of each task were analyzed separately. The data were modeled at the individual level with regressors for each condition (corresponding task and control conditions of each task) using SPM8 (Wellcome Department of Imaging Neuroscience, London; www.fil.ion.ucl.ac.uk/spm). For the programming task, additional regressors were added for the response prompt and the task probe; for the arithmetic task, an additional regressor was included for the task probe; and for the reasoning task, an additional regressor was included for

the response prompt. The regressors were convolved with the canonical hemodynamic response function (HRF). A 1/128 Hz high-pass filter was applied to remove low-frequency noise, with the AR(1) model used to account for serial correlations.

To estimate the scope of activation change associated with programming, we calculated the first-level and then the group-level contrast of the programming condition versus its control condition. Both first-level and group-level analyses were conducted using SPM8 within a gray matter mask. The mask was derived from the bilateral gray matter atlas of WFU Pickatlas (<http://www.fmri.wfubmc.edu>; Advanced Neuroscience Imaging Research Core, Wake Forest University) with one-voxel 2D dilation. The activation was similarly calculated for the four benchmark tasks. The group-level activation map of each task was thresholded by a voxel-level threshold at $P = 10^{-3}$ and by a cluster-level extent threshold based on Alphasim at $P = 0.05$. Dice coefficients, the ratio of twice the number of overlapping voxels from two different thresholded maps divided by the sum of the total number of voxels in the two thresholded maps, were calculated between the thresholded activation map of programming with those of each benchmark task to illustrate the corresponding extent of activation overlap.

Multiple Regression of Activation Pattern

Besides the univariate analysis of activation, we compared the activation patterns in the programming task and the benchmark tasks with multiple regression in a searchlight manner. Particularly, for the programming task and for each of the three benchmark tasks, we extracted the voxel-wise activation pattern of each task from the t maps of the corresponding task versus control contrasts in a given searchlight cube (5 × 5 × 5 voxels), transformed the activation pattern into a vector, and used the vector of the programming task as the dependent variable, and those of the arithmetic, the Raven reasoning, and the SD tasks as independent variables (Fig. 3). Then, we examined the resultant regression significance as well as the standardized regression coefficients of each task on the coordinate of the central voxel of the searchlight cube. In this way, by moving the searchlight cube around, we generated a map of cortical regions where the local activation pattern in the programming task can be significantly predicted by that of the three benchmark tasks and the maps of standardized regression coefficients for each regressor which indicate the unique contribution of each benchmark task in predicting the local activation pattern of the programming task. The searchlight analysis was constrained within the gray-matter mask, and cubes with less than 80 voxels within the gray matter mask were excluded. Each map was further thresholded with FDR correction for multiple comparison with $q = 0.05$. We also calculated the Dice coefficients between the pattern resemblance maps of the three benchmark tasks to illustrate the corresponding extent of overlap. We did not include the language task in the multiple regression analysis, as well as the activation and regression profile analyses (see below), simply because it showed little activation overlap with the programming task in the brain.

Analysis of Activation Profile and Regression Profile

To investigate the functional role of different brain regions activated during programming, we conducted voxel-wise clustering of activation and regression profiles across all the voxels ($n = 756$) within the thresholded activation map of the programming task

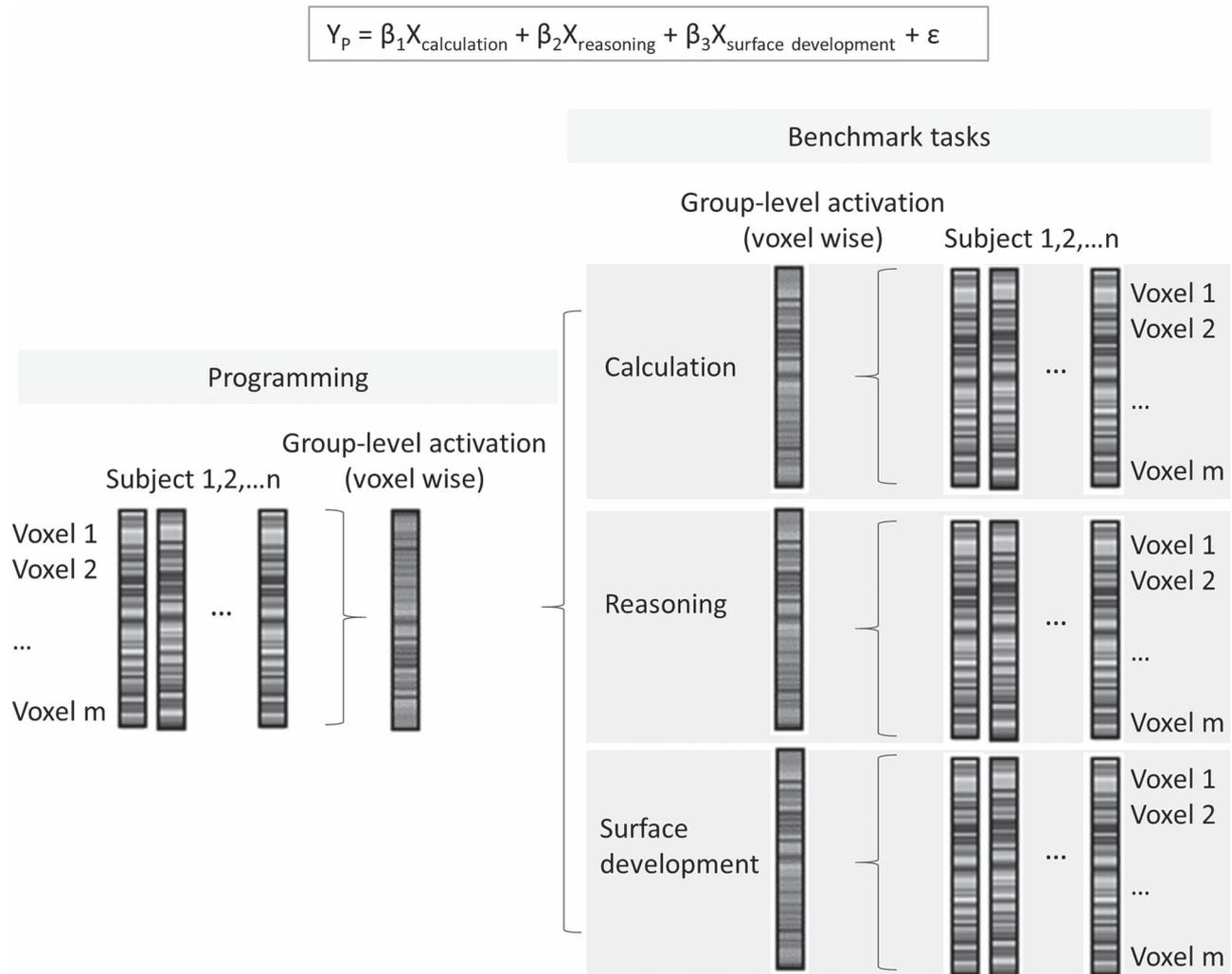


Figure 3. Illustration of the multiple regression analysis of local activation pattern.

and with their regression coefficients estimated in the multiple regression analysis. Across these voxels, we standardized the activation value for each task included in the multiple regression analysis, formed a standardized activation profile vector for each voxel, then concatenated its standardized activation profile and its regression profile (consisting of the standardized regression coefficients of each benchmark tasks), and subjected the resultant activation-regression vectors of all the voxels into k -means clustering with $k=4$, the distance metric being the squared Euclidean distance between the activation-regression vectors. The clustering was repeated five times with new initial cluster centroid positions, and the reported solution was the one with the lowest within-cluster sums of point to centroid distances.

The k value was chosen by a cross-validation approach based on stability analysis (Lange et al. 2004) before the clustering analysis. We divided the voxels randomly into a training group and a validation group, did clustering analysis with each group separately with a range of k value between 2 and 8, examined the stability between the clustering analysis of the same k values, and repeated such random division and stability calculation for 40 times. For each repetition, the validation voxels will be assigned to the training-group cluster with a centroid nearest to

this voxel, and the (dis)agreement of this assignment with the assignment based on validation-group clustering was indicated by the proportion of voxels of mismatched assignments. The instability for each k value can be illustrated by the average (dis)agreement across repetitions divided by the chance level (dis)agreement of the same k value. The k value minimizing the instability index was chosen for the formal clustering analysis including all the voxels.

Functional Connectivity Analysis

To examine the potential functional dissociation between the components identified in the k -means analysis, we examined the functional connectivity among the components (we term these brain regions collectively as the programming network) and that between each of the components with regions out of this network and compared the effect of task state on them during the programming task. More specifically, we used the four k -means components as seed ROIs. For each of them, we extracted and averaged the corresponding seed-to-voxel functional connectivity with the voxels in the other three seed ROIs as this ROI's In connectivity, and the seed-to-voxel functional connectivity with all the gray-matter voxels not

belonging to any seed ROIs as this ROI's Out connectivity. Task-dependent functional connectivity was estimated for each participant using the generalized psychophysiological interaction (gPPI) method (McLaren et al. 2012) using CONN toolbox (Whitfield-Gabrieli and Nieto-Castanon 2012), an open-source MATLAB-based cross-platform software for the analysis of functional connectivity, to measure the task-induced connectivity differences without the impact from the task-activation effects and non-task-specific correlations. In gPPI analysis, the general linear model (GLM) contains condition-specific task regressors, "seed" time-course regressors, and separate condition-specific interaction regressors. The interaction regressors predict correlation with the seed region signal during time points belonging to specific experimental condition and were used as the index of the corresponding functional connectivity. Our GLMs included condition-specific regressors for the programming condition and its control condition in addition to the task regressors of no interest for fixation, response, and response prompt presentation. We submitted the In and Out connectivity of each *k*-means component in the two experimental conditions of interest into a repeated-measure ANOVA with seed component, task state (programming vs. control), and network (In vs. Out) as independent factors. We were particularly interested in the interaction between network and task and whether this interaction, if significant, was modulated by the seed component.

Results

Neural Correlates of CT and Its Relation with Benchmark Cognitive Processes

To identify the regions involved in CT, we calculated the group-level contrast of the programming condition versus the control condition within the gray matter mask. This contrast revealed a left-lateralized network of CT-related regions in the lateral and medial cortex of the brain (Fig. 4a), spanning from the parietal to the frontal lobes ($P_{\text{uncorrected}} < 0.001$, corrected by Alphasim at $P = 0.05$). Specifically, in the parietal lobe, we found left-lateralized activation at the inferior parietal lobule, and in the frontal lobe, activation can be observed covering the bilateral middle frontal gyrus, the bilateral superior frontal gyrus, the medial part of the left superior frontal gyrus, and the triangular part of the left inferior frontal gyrus. See Table 1 for a full list.

The group-level activations were calculated and thresholded following the same approach for the four benchmark tasks ($P_{\text{uncorrected}} < 0.001$, corrected by Alphasim at $P = 0.05$). Figure 4b shows that there was little overlap of activation between the programming and the language processing, except there were five voxels in the left inferior parietal lobule. By contrast, there was substantial overlap of activation between programming and each of the three benchmark tasks (arithmetic, Raven reasoning, and SD), respectively (Fig. 4c). Visual inspection reveals that the overlap of activation between programming and these three tasks was most noticeable in the bilateral inferior parietal cortex, the left middle and superior frontal gyri as well as the medial part of the left superior frontal gyrus. Table 2 shows the activation overlap between the programming task and each of the benchmark tasks. Dice coefficients further showed that the overlap was the largest with Raven Reasoning and the smallest with the language task (Fig. 4d). Besides, there were clusters activated only in the programming task but not in the benchmark tasks, which were mainly located in the bilateral middle and superior

frontal gyri, the bilateral inferior parietal lobule, and the caudate (Table 3).

One may argue that univariate activation might not be sensitive to certain functional dissociation between tasks as different processes may activate the same region (e.g., Duncan 2010); therefore, we further used multivariate pattern analysis to examine the relation of programming to the benchmark tasks. To do this, we extracted the voxel-wise activation pattern from the searchlight cubes ($5 \times 5 \times 5$ voxels) in the group-level *t* map of the programming task and each of the three benchmark tasks, except the language task that showed little overlap with the programming task. Then, the activation pattern of the programming task was treated as the dependent variable and those of the three tasks as independent variables. The regression analysis revealed extensive cortical regions where the activation pattern of the programming task can be predicted by the activation patterns of at least one benchmark task within the regions activated by the programming task (Fig. 5b). Similar to the univariate activation analysis, the region where the coefficients of a specific benchmark task reached significance substantially overlapped, as reflected by the Venn diagram (Fig. 5c) and Dice coefficients (Fig. 5d). These regions include parts of the bilateral inferior parietal lobule, the bilateral superior and middle frontal gyri, and the left inferior frontal gyrus (Table 4). Taken together, programming recruited cortical regions involved in arithmetic, reasoning, and spatial cognition.

Component Processes Underlying Programming

To further disentangle the functional involvement of different cortical regions activated during CT, here, we examined the functional properties of each CT-related cortical region during the programming task. Because the univariate and multivariate analyses revealed shared but also separate regions, both activation magnitude and pattern were included in the clustering analysis. Specifically, we conducted a voxel-wise *k*-means clustering analysis to explore the underlining functional components shared across clusters by combining the activation profile and the pattern profile. For each voxel in the intersection of the activation map of the programming task and the regression map shown in Figures 4a and 5, we concatenated its standardized activation profile and the standardized regression coefficients and subjected the resultant activation-regression vectors into *k*-means clustering with $k = 4$, which was chosen by a cross-validation approach based on stability analysis (Lange et al. 2004, see Figure 6a for the results of the stability analysis). Among the four resultant components, the first one was mostly observed in the left medial frontal regions extending to the middle cingulate gyrus, the left premotor, the left middle frontal gyrus in and around BA 46 as well as small clusters in the left posterior parietal cortex. This component showed relatively high activation during programming and all the three nonverbal benchmark tasks, therefore was termed as the "Shared" component. Besides, this component also showed relatively high pattern similarity between programming and the spatial cognition task. The second component showed high pattern similarity between the programming and the arithmetic tasks, and was therefore coined as the "Calculation" component. The voxels belonging to this component mostly exist in the bilateral inferior parietal sulcus and in the bilateral dorsal lateral prefrontal cortex in and around BA 6 and BA 8. Besides, this component also showed overall high activation in all three benchmark tasks. The third component showed selective activation and pattern similarity

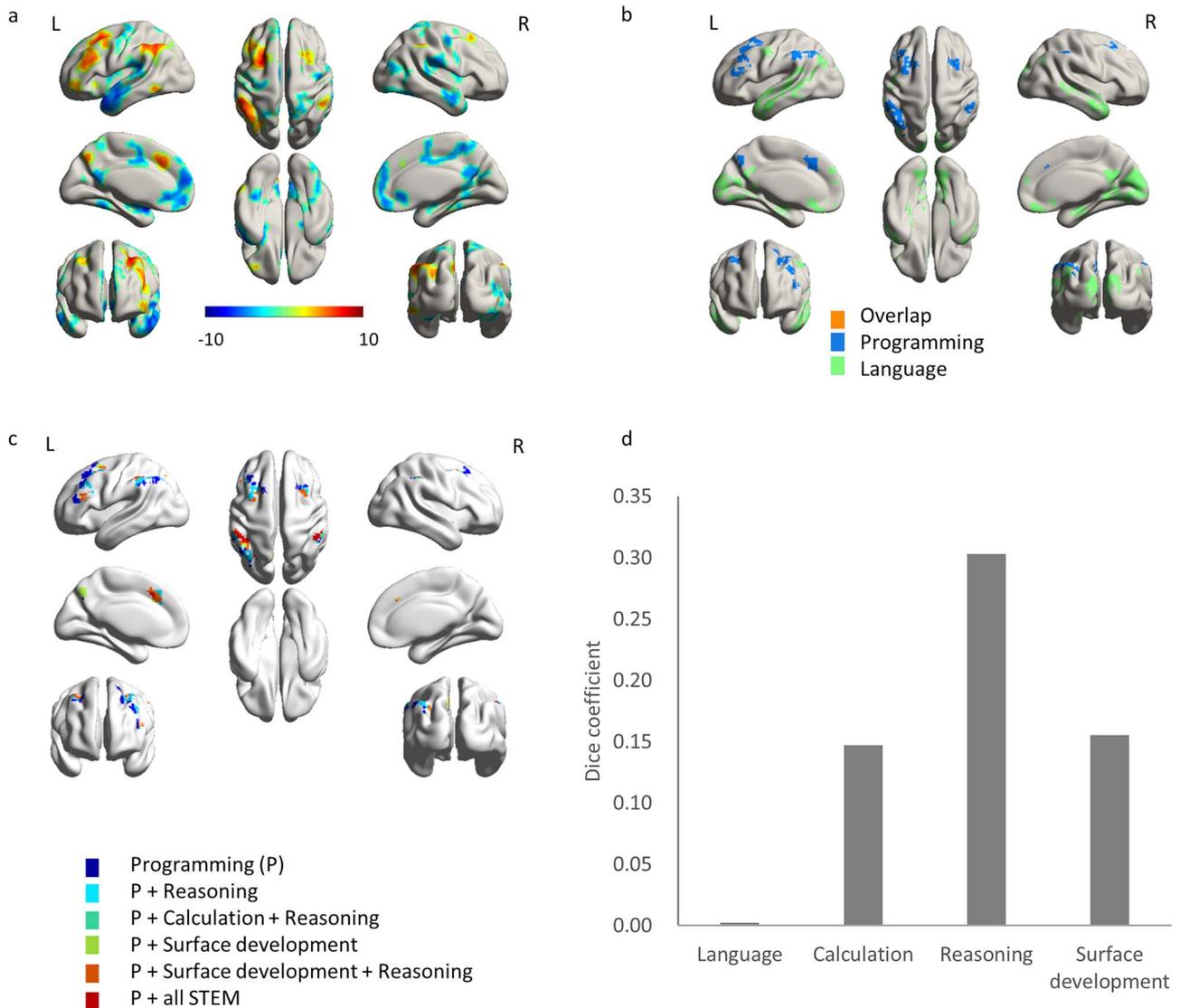


Figure 4. (a) Group-level activation map of programming. The colors denote the t values of the group level contrast between programming and its control condition. (b) Overlap of group-level activation between programming and language. (c) Overlap of group-level activation between programming and the rest of the benchmark tasks, including reasoning, arithmetic, and spatial cognition. (d) The Dice coefficient between each of the benchmark tasks with programming. The activation maps were thresholded with $P_{\text{uncorrected}} < 0.001$ and Alphasim corrected at $P = 0.05$.

in programming and reasoning, and was therefore labeled as a “Reasoning” component. The voxels belonging to this component can be found in the left middle and superior frontal gyri and the left inferior parietal lobule. The fourth is a “Visuospatial” component because it showed high pattern similarity between the programming and the spatial cognition tasks as well as high selective activation in the spatial cognition task. Voxels belonging to this component were observed in the bilateral frontal regions, and this component dominated the caudate and the left precuneus clusters activated during programming.

Another illustration of the dissociation between CT’s components is their neighboring but dissociable presence in two high-level cortical regions, the left lateral frontal cortex and the left inferior parietal cortex (Fig. 7). The left lateral frontal cortex contains voxels of all four components, with the Visuospatial component observed in the superior frontal gyrus in BA 6 and in the left inferior frontal gyrus in BA 46; the Reasoning

component can be observed in the frontal middle gyrus and the ventral part of the superior frontal gyrus, between the two patches of the Visuospatial component in this region. Voxels of the Calculation component and the major part of a cluster of the Shared component can be observed posterior to the Reasoning component in the superior frontal sulcus, near its conjunction with the precentral sulcus. Another cluster of the Shared component was observed posterior to the ventral cluster of the Visuospatial component in BA 46. The Visuospatial and the Shared components were also observed in the right superior frontal gyrus. Another cortical region occupied by multiple components is the left inferior parietal cortex, where the activation during programming can be observed along the left intraparietal sulcus, with the Calculation component taking the lower bank of the intraparietal sulcus. Ventral to it are voxels of the Shared component and of the Reasoning component, extending to the angular gyrus. Within the right inferior parietal regions activated

Table 1 Regions showing increased activation during the programming versus the control condition, $P_{\text{uncorrected}} < 0.001$, Alphasim corrected at $P = 0.05$

Hemisphere	Region	MNI coordinates			Z score	Volume (mm ³)
		x	y	z		
Left	The middle frontal gyrus	-24	15	48	5.72	4833
	The superior frontal gyrus, medial	-9	24	39	4.69	3672
	The superior frontal gyrus	-21	15	48	4.97	1728
	The inferior parietal lobule	-51	-39	45	4.91	9045
	The inferior frontal gyrus	-45	39	21	4.67	1701
	The caudate	-15	12	9	4.01	1296
	The precuneus	-9	-66	45	4.18	945
	The inferior parietal lobule	51	-42	45	4.15	1512
Right	The right middle frontal gyrus	30	18	45	3.55	1350
	The right superior frontal gyrus	24	24	57	3.94	486

Table 2 The overlap of activation between the programming task and each of the benchmark tasks, $P_{\text{uncorrected}} < 0.001$, Alphasim corrected at $P = 0.05$

Benchmark task	Region	MNI coordinates			Z score	Volume (mm ³)	
		x	y	z			
Language	L	The inferior parietal lobule	-36	-75	39	3.44	135
Arithmetic	L	The inferior parietal lobule	-42	-42	36	4.41	1458
		The precuneus	-15	0	18	3.93	162
		The superior frontal gyrus, medial	0	21	42	3.53	567
Raven reasoning	R	The right middle frontal gyrus	51	-39	48	4.03	810
	L	The superior frontal gyrus, medial	-6	30	36	5.41	2403
		The middle frontal gyrus	-27	9	48	5.10	3024
		The inferior parietal lobule	-36	-66	42	5.04	6129
		The inferior frontal gyrus	-48	18	30	4.91	891
		The caudate	-12	9	6	4.03	162
		The superior frontal gyrus	-21	6	48	3.72	351
		The precuneus	-33	3	48	3.55	54
R	The inferior parietal lobule	45	-42	45	5.04	1053	
SD	L	The right middle frontal gyrus	39	15	54	4.78	864
		The inferior parietal lobule	-39	-42	39	5.93	3699
		The middle frontal gyrus	-24	3	48	5.45	1350
		The superior frontal gyrus, medial	-3	18	39	5.11	1701
		The superior frontal gyrus	-21	3	57	4.84	459
		The caudate	-12	3	9	4.73	162
		The inferior frontal gyrus	-51	30	21	4.24	945
		The precuneus	-33	3	48	3.76	27
		R	The inferior parietal lobule	45	-42	45	4.82
	The right middle frontal gyrus		30	12	51	4.38	540

during programming, all the voxels fell into the Calculation component (Fig. 7).

The four components suggested division of labor underlying CT. Therefore, a critical question is how they collaborate during programming. To address this question, we used the four components as seed ROIs and then calculated the functional connectivity among the seed ROIs. The averaged functional

connectivity among seed ROIs is the measure for the integration of the programming network constructed by these seed ROIs (i.e., In connectivity). By contrast, the functional connectivity among the seed ROIs and voxels that were not activated by the programming task is the measure for the separation between the programming network and the rest of the brain (i.e., Out connectivity) (Wang et al. 2016; Hao et al. 2018; Yu et al. 2018).

Table 3 Regions showed activation only in programming but not in any of the benchmark tasks, $P_{\text{uncorrected}} < 0.001$, Alphasim corrected at $P = 0.05$

Hemisphere	Region	MNI coordinates			Z score	Volume (mm ³)
		x	y	z		
Left	The middle frontal gyrus	-21	21	45	4.61	2052
	The superior frontal gyrus	-21	18	48	4.56	1107
	The inferior parietal lobule	-48	-57	45	4.37	1782
	The inferior frontal gyrus	-45	36	15	4.26	189
	The caudate	-15	12	39	3.84	756
	The superior frontal gyrus, medial	-12	24	33	3.01	81
	Right	The right superior frontal gyrus	24	24	57	3.77
The right middle frontal gyrus		27	21	54	3.41	486
The inferior parietal lobule		54	-45	48	3.07	162

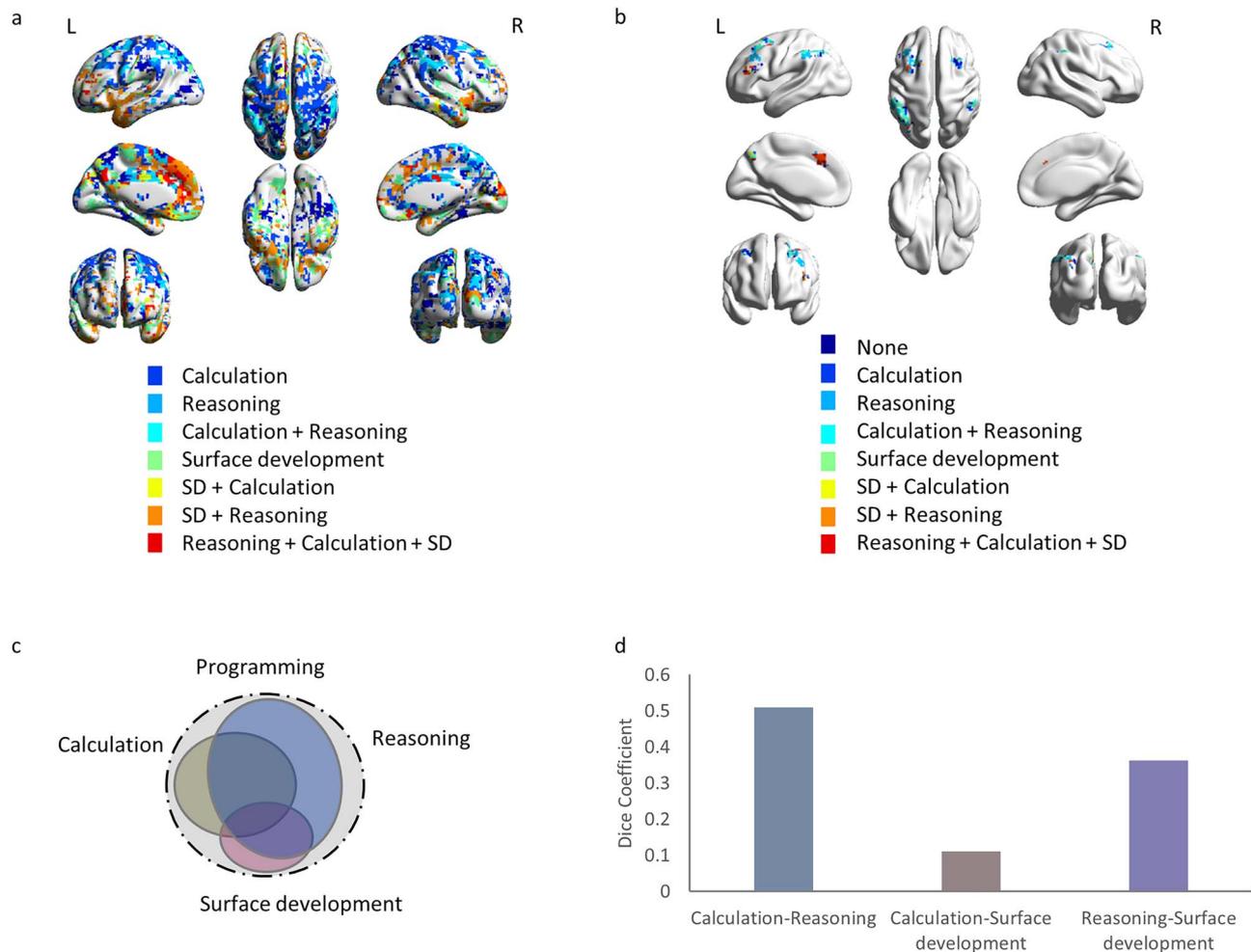


Figure 5. (a) Regions where activation patterns in the arithmetic, reasoning, or spatial cognition task significantly predicted that of the programming task. (b) Pattern resemblance within regions showing activation increase in programming. (c) The dissociation between the pattern resemblance of the three benchmark tasks. (d) The Dice coefficients between the pattern resemblance maps of the three benchmark tasks.

Table 4 Cortical regions showing pattern resemblance between benchmark tasks and programming

Benchmark task	Region	Region	MNI coordinates ^a			Peak β^b	Volume (mm ³) ^c
			x	y	z		
Arithmetic	L	The inferior parietal lobule	-48	-42	57	0.90	4320
		The superior frontal gyrus	-21	12	63	0.80	756
		The middle frontal gyrus	-24	12	57	0.58	1485
	R	The superior frontal gyrus, medial	-15	12	63	0.55	594
		The caudate	-12	0	21	0.40	135
		The inferior frontal gyrus	-45	45	12	0.37	54
Raven Reasoning	L	The inferior parietal lobule	51	-45	51	1.00	999
		The right middle frontal gyrus	33	12	57	0.65	405
		The right superior frontal gyrus	24	24	57	0.48	243
	R	The inferior parietal lobule	-60	-42	39	1.12	6372
		The middle frontal gyrus	-39	36	33	1.06	4428
		The superior frontal gyrus, medial	-3	27	33	0.84	2646
Spatial cognition	L	The superior frontal gyrus	-21	15	51	0.83	702
		The inferior frontal gyrus	-51	24	27	0.78	702
		The precuneus	-36	6	48	0.70	54
	R	The caudate	-15	6	15	0.55	783
		The right middle frontal gyrus	33	27	51	0.73	540
		The right superior frontal gyrus	27	27	54	0.67	378
Multiple tasks	L	The inferior parietal lobule	57	-36	51	0.64	486
		The superior frontal gyrus, medial	-12	24	45	0.80	324
		The middle frontal gyrus	-42	42	21	0.75	567
	R	The inferior frontal gyrus	-45	33	15	0.65	729
		The superior frontal gyrus, medial	-9	24	45	0.62	2295
		The inferior parietal lobule	-33	-69	39	0.58	297
All	L	The right middle frontal gyrus	33	12	48	0.36	162
		The inferior parietal lobule	-33	-78	36		3834
		The superior frontal gyrus, medial	-6	33	30		2403
	R	The middle frontal gyrus	-42	39	15		1836
		The inferior frontal gyrus	-48	42	12		594
		The superior frontal gyrus	-12	24	39		459
None	L	The caudate	-15	3	15		81
		The inferior parietal lobule	54	-45	45		486
		The right superior frontal gyrus	24	18	48		243
	R	The right middle frontal gyrus	33	24	48		189
		The superior frontal gyrus, medial	-6	33	30		432
		The superior frontal gyrus	-42	45	12		54
None	L	The middle frontal gyrus	-45	45	15		27
		The inferior parietal lobule	-36	-45	27		1215
		The middle frontal gyrus	-42	45	15		729
	R	The superior frontal gyrus	-18	21	42		432
		The inferior frontal gyrus	-42	33	15		378
		The caudate	-12	9	6		270
None	R	The superior frontal gyrus, medial	-9	30	30		54
		The right middle frontal gyrus	30	15	45		432
		The inferior parietal lobule	48	-48	39		216
None	R	The right superior frontal gyrus	21	21	51		108

Note:^aFor each benchmark task, this column reported the center coordinates of the searchlight cube in which the standardized regression coefficients of the given task reached its maximum in a given anatomical region. For the rest, this column reported the center coordinates of one example searchlight cube which fulfilled the given condition.

^bThe peak standardized regression coefficient in a given anatomical region.

^cThe volume of voxels centered on which the searchlight cubes fulfilled the given condition.

We submitted the In and Out connectivity of each component into a repeated-measure ANOVA with component (Shared, Visuospatial, Calculation, and Reasoning), task state (programming vs. control), and network (In vs. Out) as independent variables. The results revealed a significant main effect of component, $F(1,19)=4.32$, $P=0.008$, partial $\eta^2=0.19$, with connectivity with

the Shared component being generally lower than that with the Calculation component ($MD=0.004$, $P=0.029$) while no other pairwise difference was significant. There was also a main effect of task state, $F(1,19)=17.94$, $P<0.001$, partial $\eta^2=0.49$, with functional connectivity being higher during programming than the control condition. Critically, there was a significant

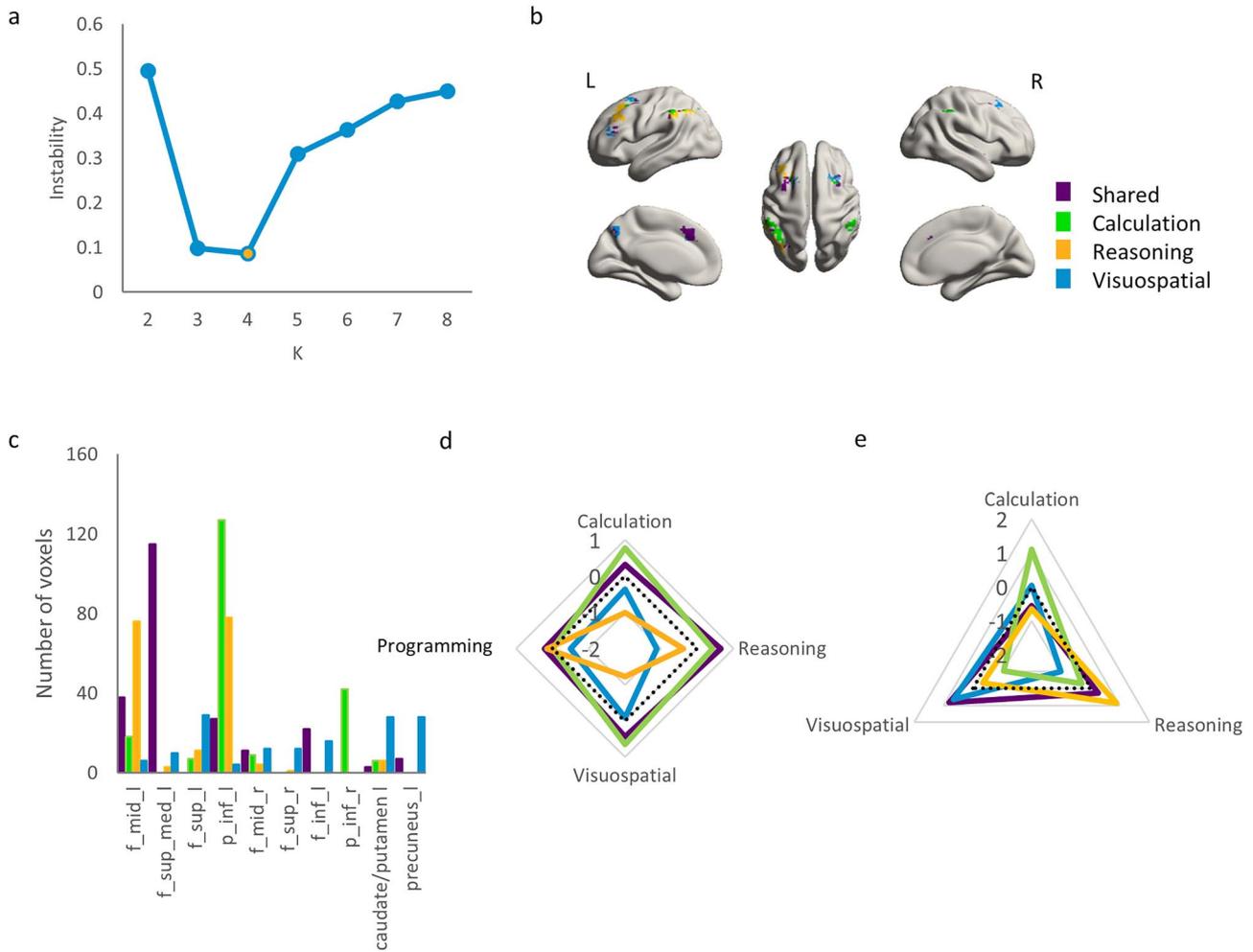


Figure 6. The results of the *k*-means clustering analysis revealed cognitive components across anatomical regions. (a) *k* = 4 was selected by a cross-validation approach based on stability analysis (Lange et al. 2004). (b) The spatial extents of the four resultant components. (c) The composition of the four components in each anatomical cluster. (d) The standardized activation profile of the four components. The activation level during programming was not entered into the clustering analysis. It was presented here as a benchmark. (e) The profile of the standardized regression coefficients of the four components.

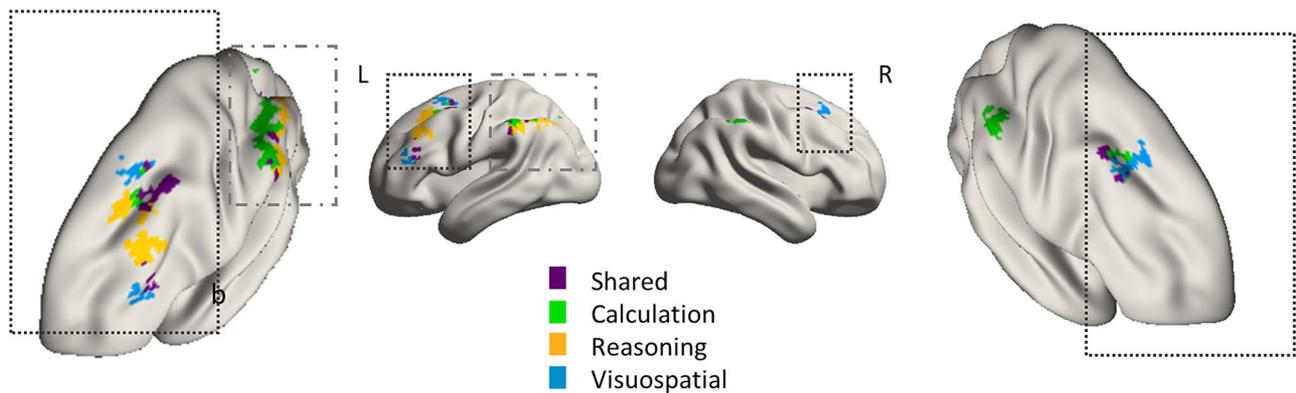


Figure 7. The location of the four components in the bilateral lateral frontal cortex (solid-line frames) and the left inferior parietal cortex (dash-line frames).

interaction between task state and network, $F(1,19)=37.08$, $P<0.001$, with the In connectivity being higher ($MD=0.007$, $P<0.001$) and the Out connectivity being lower ($MD=0.002$, $P=0.001$) during programming than in the control condition

(Fig. 8), suggesting that the functional connectivity among components were strengthened during the programming task. This interaction was not modulated by component and the three-way interaction was not significant ($P=0.468$).

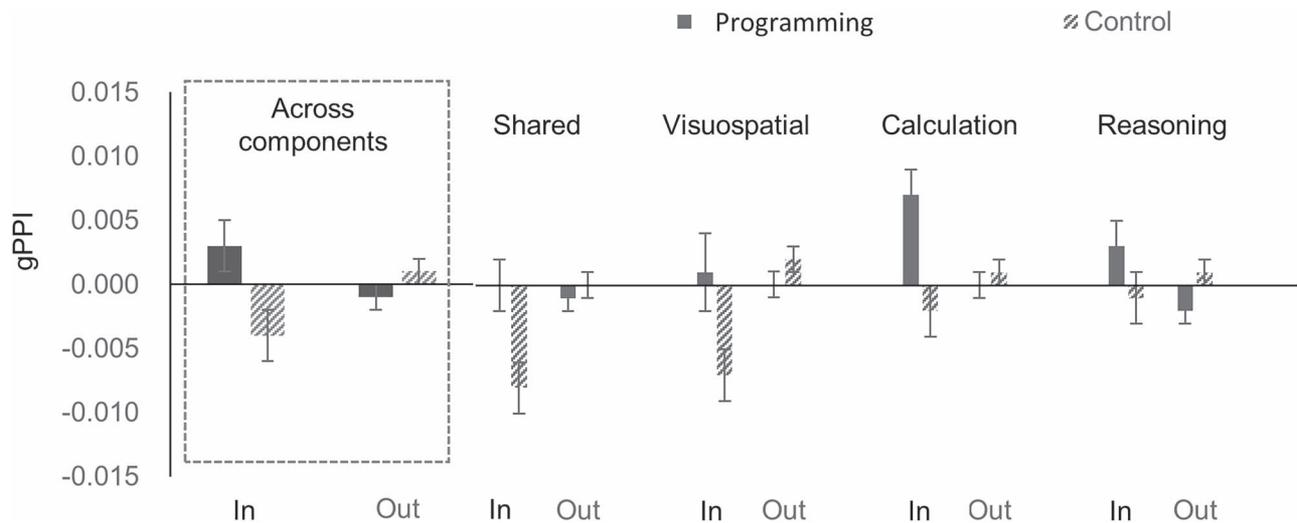


Figure 8. The functional connectivity among the components of the programming network (the In connectivity) and between the components and regions not responsive during programming (the Out connectivity). The bars in the dashed-line frame illustrated the In and Out connectivity across components, which shows that the In connectivity increased during programming comparing with the control ROI, while the Out connectivity decreased. The bars right to the dashed-line frame illustrated the In and Out connectivity with each component as the seed ROI, which all showed the same pattern.

Discussion

The present study used fMRI to explore the neural basis of CT and to further identify its underlying cognitive components. We found that regions involved in CT were left-lateralized, mainly in the bilateral inferior parietal lobule, the bilateral middle and superior frontal gyri, the medial part of the left superior frontal gyrus, and the triangular part of the left inferior frontal gyrus. By comparing the neural correlates of CT and a set of key cognitive tasks, namely language, arithmetic, reasoning, and spatial cognition, we observed that CT showed extensive similarity with all but the language task in both activation location and local activation pattern, suggesting that, instead of a unitary capacity, CT is built upon multiple cognitive processes, requiring a composite set of abilities. Further, a *k*-means clustering analysis revealed a four-component organization underlying CT, and the connectivity analysis revealed online collaboration among these dissociable components. The present study provides one of the first empirical evidence on the neural basis of CT and further sheds light on the application of CT in practice.

Similar to previous studies, the present study observed the involvement of distributed fronto-parietal regions in coding-related tasks. The activation of the inferior and the middle frontal gyri and the inferior parietal lobule has been observed during code comprehension (Siegmund et al. 2014; Castelhana et al. 2019), and the blood oxygen level-dependent (BOLD) signal from a set of distributed brain regions in frontal, parietal, and occipitotemporal regions can readily classify code reviewing and code comprehension from prose reviewing (Floyd et al. 2017). However, unlike studies on code comprehension, we did not observe activation change in the insula, and the activation change in the medial frontal cortex we observed was more posterior than that modulating insula during code debugging (Duraes et al. 2016; Castelhana et al. 2019). This is probably because the programming task of the present study focused on the problem-solving, or the “production”, aspect of CT. Future work is needed to directly investigate the potential dissociation between the speculated “production” and “comprehension” distinction in CT. Having said this, we would also like

to underline that although programming differed from code review and comprehension in visual input, task format, and required output, they nevertheless shared some brain regions. That is, programming may involve at least some of the cognitive processes recruited in code review and comprehension. For instance, they may all require the recruitment of CT in terms of using computational concepts and skills to solve problems. This commonness supported the theoretical conception of CT as a composite cognitive process shared by a variety of tasks, instead of a transient collection of cognitive processes that happened to be recruited at the same time for the ongoing experimental task, and it might be part of the critical component behind various programming-related tasks regardless of their varied formats.

One major insight drawn from the present results was that instead of a unitary capacity, CT is built upon multiple key cognitive processes and requires a composite set of abilities. First, among the four benchmark tasks, namely, the language, the arithmetic, the reasoning, and the spatial cognition tasks, all but the language task showed extensive similarity with CT in activation location and pattern. Further, an additional contrast of task difficulty (Supplementary Analysis 1) revealed that the neural activation of the mental programming task can be hardly explained by the domain-general processes dealing with task difficulty, with extensive regions activated in the CT task not being sensitive to task difficulty. Instead, CT is characterized by the division of labor among the cortical regions. Though most regions activated in CT were also activated in at least one of the benchmark tasks, their activation and pattern similarity profile across tasks differed. Clustering analysis revealed a four-component organization among the cortical regions activated during programming. Among the four components, three can be reasonably speculated to reflect cognitive functions that were differentially associated with different cognitive processes.

The first is a Reasoning component, which was named because its high activation and high level of activation pattern similarity between programming and the reasoning task suggested its involvement in reasoning-related processes. The emergence of a Reasoning component is consistent with the

previous finding of psychometrical studies that the performance in a test of CT is correlated with reasoning and problem-solving abilities (Roman-Gonzalez et al. 2017). Voxels of this component can be found in the left middle and superior frontal gyri and the left inferior parietal lobule. Particularly, the left dorsal lateral prefrontal cluster of the Reasoning component partially corresponds to de la Vega et al.'s (2018) cluster 9 and C6 identified based on the functional connectivity by Goulas et al. (2012). This region has strong functional connectivity with the default mode network (Goulas et al. 2012) and may be involved in "internal" processes such as introspection during programming and recollection of previous thoughts in solving a problem. The left inferior parietal lobule voxels of the Reasoning component were largely in the ventral part of the inferior parietal lobule, falling in the parietal regions of the default mode network (Shulman et al. 1997). In addition, this cluster extends into the angular gyrus. The angular gyrus has been proposed to be an interface between executive processes and stored episodic representations (Wagner et al. 2005; Vilberg and Rugg 2008), which might be particularly typical in solving complex problems as in programming and reasoning. In addition, due to its partial overlap with the default mode network, we speculate that this component reflects the divergent thinking process in programming, or the generation phase of problem-solving (Kleinmintz et al. 2019). Previous studies proposed that the default mode network might be involved in the generative cognitive processes rather than the error detection and evaluation phase in creative thinking (Beatty et al. 2016; Kleinmintz et al. 2019). Consistent with this speculation, the Reasoning component showed selective activation in the reasoning and the programming tasks, the only two tasks in our battery that may involve divergent thinking.

The second is a Calculation component because the voxels of this component showed high activation and high level of activation pattern similarity with programming in the arithmetic task. The voxels of this component are mainly observed in the bilateral intraparietal sulcus and in the bilateral dorsal lateral prefrontal cortex in and around BA 6 and BA 8. The emergence of a Calculation component is consistent with the existing theoretical proposal that computer programming is associated with STEM abilities (Fedorenko et al. 2019). We speculate that this component reflects the quantity representation during programming. Consistent with this speculation, the bilateral inferior parietal lobule, especially the intraparietal sulcus, has been established to be critical for mathematical cognition. They were activated during calculation tasks (Davis et al. 2009; Fedorenko et al. 2013; Humphreys and Lambon Ralph 2015) and are proposed to be associated with quantity representation (Dehaene et al. 2003; Arsalidou and Taylor 2011), the processing of abstract mathematical formulae (Friedrich and Friederici 2009), and algebraic transformation (Anderson et al. 2003). The bilateral dorsal lateral prefrontal voxels of this component spatially correspond to the anterior part of de la Vega et al.'s (2018) cluster 6/8. Activation in this region has been reported in arithmetical computation tasks, probably to support the phonological processing involved in numerical tasks (Zhou et al. 2018). However, an additional contrast of task difficulty (Supplementary Analysis 1) revealed that some of the parietal voxels of this component were sensitive to task difficulty, which is consistent with the observation that the voxels in this component showed an overall high activation in all three benchmark tasks. This finding suggested that even with selective pattern similarity profile, the Calculation component might not be entirely domain-specific

as its name suggested, and future work is needed to explore the functional specificity of these voxels in dealing with arithmetic problems.

The third is a Visuospatial component because the voxels of this component showed high activation and high level of activation pattern similarity between programming and the spatial cognition task, suggesting their involvement in visuospatial processes. The emergence of a Visuospatial component is consistent with the previous finding from psychometrical studies that the performance in a test of CT, a cognitive ability considered behind programming, is correlated with visuospatial abilities (Roman-Gonzalez et al. 2017). The voxels of this component were mainly observed in the left caudate, the left precuneus, the bilateral superior frontal gyrus in BA 6, and the left inferior frontal gyrus in BA 46. Activation in the posterior parietal lobe, including the left precuneus, has been observed in visuospatial tasks involving both small- and large-scale spatial abilities (Li et al. 2019). The bilateral superior frontal gyrus in BA 6 may correspond with the cluster 6/8 identified by meta-analysis based on whole-brain co-activation, which is proposed to be involved in attention and eye movement (de la Vega et al. 2018), and has been found involved in mental imagery (Winlove et al. 2018). The voxels in the left inferior frontal gyrus in BA 46 may belong to the rostral part of the cluster 9/46v or the posterior part of the cluster 9/46v identified by meta-analysis based on whole-brain co-activation, which is involved in working memory and executive control (de la Vega et al. 2018). The caudate is part of the basal ganglia and is involved in navigation tasks (Latini-Corazzini et al. 2010). Taking the known functions of these regions into account, we speculate that this component reflects the visuospatial representation and mental imagery processes involved in programming.

Even though the clustering analysis suggested functional dissociation during programming, the fourth component, the Shared component, emerged besides these three, at least in terms of activation and pattern similarity, relatively domain-specific components. This component and the findings of the connectivity analysis suggested that CT may not be a mere co-activation of parallel cognitive processes, but a composite process integrated as a whole. Voxels of the Shared component showed high activation in the programming task as well as in all the three nonverbal benchmark tasks and showed high level of activation pattern similarity between programming and the spatial cognition task. These voxels are predominantly observed in the left premotor cortex, the left medial frontal regions extending to the middle cingulate gyrus, the left middle frontal gyrus in and around BA 46 as well as small clusters in the left posterior parietal cortex. This component remarkably overlaps with the attention and working memory system (Cabeza and Nyberg 2000) in the ventrolateral prefrontal region in and around BA 46, the premotor region, and the inferior parietal lobule. In the premotor region and the inferior parietal lobule, this component also overlaps with the cognitive control network (Cole and Schneider 2007), the working memory "core" network (Rottschy et al. 2012), the dorsal attention network (Corbetta and Shulman 2002; Yeo et al. 2011), and the multiple demand network (Duncan 2010), which mediates goal-directed behavior and subserves the control of cognitive operation through successive task steps. Particularly, the bilateral premotor cortex was thought to be involved in the maintenance of visuospatial attention (Owen et al. 2005). It is found responding to various cognitive control demands and was considered among the "core" cognitive control regions (Dosenbach et al. 2006; Cole and Schneider

2007) and the core regions of the dorsal attention system (Vincent et al. 2008). The cingulate gyri have been related to error monitoring (Taylor et al. 2007), integration of information (DeVue et al. 2007), coordinating and integrating activity of multiple attentional systems (Peterson et al. 1999), and setting goals by integrating available information (Arsalidou and Taylor 2011). The rostral lateral prefrontal area in and around BA 46 underlies the deployment of general cognitive resources and is involved in stimulus-driven attention and working memory (Christoff and Gabrieli 2000; Curtis and D'Esposito 2003; Owen et al. 2005; Rottschy et al. 2012; de la Vega et al. 2018). Also, this region is considered relevant to fluid intelligence (Clark et al. 2017). The posterior parietal cortex was proposed mediating shifts in attention, retaining task-related temporal information, and preparing for a given task (Owen et al. 2005). The emergence of a component implicated in executive control and working memory is also consistent with the behavioral correlation between code review performance and working memory (Baum et al. 2019) and the expertise-dependent activation difference during code comprehension in the power of alpha and theta band (Crk et al. 2016; Yeh et al. 2017), both reflecting workload and working memory (Klimesch 1999; Tesche and Karhu 2000; Raghavachari et al. 2001).

Based on these functional characteristics, we speculate that this component consists of the subset of brain regions supporting executive control and the deployment of cognitive resources according to ongoing task demand. During programming, this component may function as a local integration desk where the workflow toward the solution of a programming problem was specified, the task of each the domain-specific components assigned, commands issued, and feedback gathered. In other words, the other components may deal with problems within their respective specialty, while the Shared component may weave these separate cognitive processes into an integrated CT process. Admittedly, it is also possible that this Shared component did not reflect the active integration of functions or information, but a domain-general component present in all the benchmark tasks and the programming task, such as executive control and/or inhibitory processes. Future study is needed to examine this possibility. An additional contrast of task difficulty (Supplementary Analysis 1) revealed that many of the medial frontal voxels of the Shared component were sensitive to task difficulty, which is consistent with the finding that the voxels in this component showed an overall high activation in all three benchmark tasks. This finding suggested that this component might respond to task difficulty, which reinforces the speculated role of this component in domain-general functions. Future studies thus need to additionally consider the potential influence of task difficulty in identifying the neural correlates of programming and CT.

The existence of a Shared component hinted at a potential local site of integration during programming, while the functional connectivity analysis further revealed collaboration between the functional components on a global scale. The between-component functional connectivity increased during the programming condition comparing with the control condition, suggesting that these components joined force in the task of programming and functioned as cohesive integrity. Further, in contrast to within-network connectivity, the connectivity between the other brain regions and these components decreased during programming comparing with the control condition, suggesting that the neural activity of the programming components, during their collaboration in

programming, deviates from that of the other parts of the brain, that is, these components encapsulated programming. Note that the present study focused on examining the potential encapsulation of the CT network; therefore, in our analysis, the Out connectivity of each participant was estimated by an overall average of seed-based connectivity to all voxels in regions other than CT network, which were much more extensive and presumably more functionally varied than those within the CT network.

In summary, our results suggested that programming may involve 1) a reasoning module probably involving creativity/divergent thinking via the default mode network, 2) a quantity processing module, 3) a visuospatial processing module, and 4) a shared module probably functioning as a local integration hub via domain-general executive control and working memory. These components, however dissociable, were involved in global collaboration reflected by the connectivity analysis. Together, the global integration reflected by the connectivity analysis and the (speculated) local integration effected by the Shared component supported the construct validity of conception of the cognitive substrates of programming, and CT in general, as a relatively self-contained and dissociable cognitive process, instead of a transient collection of cognitive processes happened to be recruited at the same time for the ongoing programming problem. Note that the number of components was decided in a data-driven manner by optimizing the instability index of *k*-means clustering, and it was coincidentally identical to the number of benchmark tasks. The present study tentatively named each component according to their activation and pattern similarity profiles between programming and the benchmark tasks as well as the known functional characteristics of the brain regions associated with each component. That is, the naming and the functional inference of the components came from reverse inference and shall be interpreted with caution. In addition, the naming of the components was unavoidably affected by the selection of the benchmark tasks of the present study. Therefore, future work is needed to directly test the speculated functional role of each component.

Another intriguing observation of the present study was a clear dissociation between verbal and nonverbal tasks. Unlike nonverbal tasks, the neural activation during programming showed negligible overlap with the language task. This is not consistent with the reported association between the language abilities and programming practice or the mastering of programming language (Pepler and Warschauer 2011; Kazakoff and Bers 2014) and the fMRI evidence that code comprehension activates brain regions involved in language processing, such as BA 44 in the inferior frontal gyrus and BA 21 in the middle temporal gyrus (Siegmund et al. 2014). The inconsistency might come from the difference in task characteristics. That is, previous studies mostly focused on processes related to code comprehension and reviewing, which concerns the usage of programming language to a larger extent than the programming task used in the present study, which does not require the participants to comprehend or generate any specific code. Having said this, note that the language-related activation in the present study was rather constricted in the frontal cortex comparing with that in a study that used a subject-specific fROI approach to identify the neural correlates of language (Fedorenko et al. 2010). Specifically, language-related regions in the IFG and MFG reported by Fedorenko et al. (2010) were adjacent but posterior to the programming-related regions identified in the present study. Therefore, even with larger

language-related regions suggested by previous studies, the distinction between language processing and programming likely remained, implying that programming and language processing may underlie different mechanisms. However, caution shall be exercised because the line separating them might be thin, given their adjacency in anatomy.

The potential discrepancy between the neural correlates of language and programming may further reflect a critical feature of programming and maybe CT in general: it is a composite task requiring multiple functional modules and could be understood in multiple levels. Context-dependent neural correlates associated with the semantic and/or symbolic characteristics of programming languages or with cognitive demands of specific programming problems might be quite different from the neural correlates dictated by the “core” cognitive demands overarching various contexts, types, and stages of programming. The investigation regarding the former is doubtlessly important in understanding the cognitive and neural correlates of programming as a behavioral task, while investigation on the latter helps reveal the cognitive substrates of a general approach of thinking, that is, CT (Wing 2006, 2008) behind various programming tasks. Future studies might benefit from specifying the level of generality of their operational definition and the context dependency of their findings.

Also note that the present study examined the neural correlates of CT by comparing the neural correlates of a representative CT task with the representative tasks of other cognitive domains. This unavoidably led our primary attention to corresponding domains, such as reasoning, spatial ability, and arithmetic ability. An equally important question would be to dedicatedly examine how domain-general processes, such as working memory, decision-making, and executive control, modulate the functionality of CT. Our results provided preliminary evidence of the involvement of such domain-general processes. For instance, the prefrontal, the medial frontal, and the parietal clusters of the CT network were also recruited in decision-making, particularly in the choice-selection stage (Ernst et al. 2004), suggesting a role of decision-making in CT. This is in line with the theoretical proposal that evaluation and decision-making on solutions are critical parts of CT (CSTA and ISTE 2011; Selby and Woollard 2013). Also, as we mentioned earlier, the CT network subsumed the Shared component that was likely involved in attentional control, working memory, and executive control of complex cognitive operations. However, these domain-general functions were inferred in a data-driven manner; therefore, future studies are needed to dedicatedly examine how these domain-general processes organize and modulate CT.

A factor that might affect the generality of the findings of the present study is that the present study recruited participants of a narrow range of expertise with computational tasks. They were all college students with basic knowledge and basic experience of programming. Considering the previous report of expertise-dependent neural correlates of programming, especially in the neural distinction between code processing and language processing (Floyd et al. 2017), it might deserve further investigation regarding the expertise-related factors influencing the neural and cognitive correlates of programming.

Even with these limitations, the present study presented one of the first investigations on the neural and cognitive substrates of programming, a representative specimen of CT. We observed that programming reused some well-established cognitive functions but at the same time assembled them adaptively, for instance, by adding quantity processing and visuospatial processing into reasoning and divergent thinking, to meet the new

cognitive challenge presented by technique advance. Investigation following this line may advance the understanding of the potential of human intelligence and its interplay with culture.

Conclusion

Using fMRI, the present study identified a largely left-lateralized network of distributed cortical regions that are involved in computational thinking, mainly in the bilateral inferior parietal lobule, the bilateral middle and superior frontal gyri, the medial part of the left superior frontal gyrus, and the triangular part of the left inferior frontal gyrus. This network consists of multiple dissociable components among which evident online collaboration exists. These findings suggest that CT is not a unitary capacity or a simple sum of parallel cognitive processes, but a composite cognitive process integrating a set of intellectual abilities, particularly those in the STEM domains.

Supplementary Material

Supplementary material can be found at *Cerebral Cortex* online.

Funding

Shenzhen Institute of Artificial Intelligence and Robotics for Society (AC01202005022); the National Natural Science Foundation of China (31861143039, 31600925), the Fundamental Research Funds for the Central Universities (2019NTSS27), and the National Basic Research Program of China (2018YFC0810602).

Notes

Conflict of Interest: None declared.

References

- Amalric M, Dehaene S. 2016. Origins of the brain networks for advanced mathematics in expert mathematicians. *Proc Natl Acad Sci.* 113:4909–4917.
- Ambrósio AP, Xavier C, Georges F editors. Digital ink for cognitive assessment of computational thinking, 2014 *IEEE Frontiers in Education Conference (FIE) Proceedings*; 2014. IEEE. p. 1–7.
- Anderson JR, Qin YL, Sohn MH, Stenger VA, Carter CS. 2003. An information-processing model of the BOLD response in symbol manipulation tasks. *Psychon Bull Rev.* 10:241–261.
- Arsalidou M, Taylor MJ. 2011. Is $2 + 2 = 4$? Meta-analyses of brain areas needed for numbers and calculations. *Neuroimage.* 54:2382–2393.
- Bastin J, Deman P, David O, Gueguen M, Benis D, Minotti L, Hoffman D, Combrisson E, Kujala J, Perrone-Bertolotti M. 2016. Direct recordings from human anterior insula reveal its leading role within the error-monitoring network. *Cereb Cortex.* 27(2):1545–1557.
- Baum T, Schneider K, Bacchelli A. 2019. Associating working memory capacity and code change ordering with code review performance. *Empirical Software Engineering.* 24:1762–1798.
- Beaty RE, Benedek M, Silvia PJ, Schacter DL. 2016. Creative cognition and brain network dynamics. *Trends Cogn Sci.* 20:87–95.
- Cabeza R, Nyberg L. 2000. Neural bases of learning and memory: functional neuroimaging evidence. *Curr Opin Neurol.* 13:415–421.
- Castelhano J, Duarte IC, Ferreira C, Duraes J, Madeira H, Castelo-Branco M. 2019. The role of the insula in intuitive expert

- bug detection in computer code: an fMRI study. *Brain Imaging Behav.* 13:623–637.
- Christoff K, Gabrieli JDE. 2000. The frontopolar cortex and human cognition: evidence for a rostrocaudal hierarchical organization within the human prefrontal cortex. *Psychobiology.* 28:168–186.
- Clark CM, Lawlor-Savage L, Goghari VM. 2017. Comparing brain activations associated with working memory and fluid intelligence. *Dermatol Int.* 63:66–77.
- Cole MW, Schneider W. 2007. The cognitive control network: integrated cortical regions with dissociable functions. *Neuroimage.* 37:343–360.
- Corbetta M, Shulman GL. 2002. Control of goal-directed and stimulus-driven attention in the brain. *Nat Rev Neurosci.* 3:201–215.
- Crk I, Kluthe T, Stefik A. 2016. Understanding programming expertise: an empirical study of phasic brain wave changes. *Acm Transactions on Computer-Human Interaction.* 23:29.
- CSTA, ISTE. 2011. *Operational definition of computational thinking for K-12 education.* <https://id.iste.org/docs/ct-documents/computational-thinking-operational-definition-flyer.pdf>. Journal.
- Curtis CE, D'Esposito M. 2003. Persistent activity in the prefrontal cortex during working memory. *Trends Cogn Sci.* 7:415–423.
- Dalbey J, Linn MC. 1985. The demands and requirements of computer programming: a literature review. *J Educ Comput Res.* 1:253–274.
- Davis N, Cannistraci CJ, Rogers BP, Gatenby JC, Fuchs LS, Anderson AW, Gore JC. 2009. The neural correlates of calculation ability in children: an fMRI study. *Magn Reson Imaging.* 27:1187–1197.
- de la Vega A, Yarkoni T, Wager TD, Banich MT. 2018. Large-scale meta-analysis suggests low regional modularity in lateral frontal cortex. *Cereb Cortex.* 28:3414–3428.
- Dehaene S, Piazza M, Pinel P, Cohen L. 2003. Three parietal circuits for number processing. *Cogn Neuropsychol.* 20:487–506.
- Denning PJ. 2017. Remaining trouble spots with computational thinking. *Communications of the ACM.* 60:33–39.
- Devue C, Collette F, Baiteau E, Dequeldre C, Luxen A, Maquet P, Bredart S. 2007. Here I am: the cortical correlates of visual self-recognition. *Brain Res.* 1143:169–182.
- Dosenbach NUF, Visscher KM, Palmer ED, Miezin FM, Wenger KK, Kang HC, Burgund ED, Grimes AL, Schlaggar BL, Petersen SE. 2006. A core system for the implementation of task sets. *Neuron.* 50:799–812.
- Duncan J. 2010. The multiple-demand (MD) system of the primate brain: mental programs for intelligent behaviour. *Trends Cogn Sci.* 14:172–179.
- Duraes J, Madeira H, Castelhana J, Duarte C, Branco MC, Ieee. 2016. WAP: Understanding the brain at software debugging. In: *2016 IEEE 27th International Symposium on Software Reliability Engineering.* Ottawa, Canada, pp. 87–92.
- Ernst M, Nelson EE, McClure EB, Monk CS, Munson S, Eshel N, Zarah E, Leibenluft E, Zametkin A, Towbin K, et al. 2004. Choice selection and reward anticipation: an fMRI study. *Neuropsychologia.* 42:1585–1597.
- Fedorenko E, Duncan J, Kanwisher N. 2012. Language-selective and domain-general regions lie side by side within Broca's area. *Curr Biol.* 22:2059–2062.
- Fedorenko E, Duncan J, Kanwisher N. 2013. Broad domain generality in focal regions of frontal and parietal cortex. *Proc Natl Acad Sci USA.* 110:16616–16621.
- Fedorenko E, Hsieh P-J, Nieto-Castanon A, Whitfield-Gabrieli S, Kanwisher N. 2010. New method for fMRI investigations of language: defining ROIs functionally in individual subjects. *J Neurophysiol.* 104:1177–1194.
- Fedorenko E, Ivanova A, Dhamala R, Bers MU. 2019. The language of programming: a cognitive perspective. *Trends Cogn Sci.* 23:525–528.
- Fedorenko E, Thompson-Schill SL. 2014. Reworking the language network. *Trends Cogn Sci.* 18:120–126.
- Floyd B, Santander T, Weimer W. (2017). Decoding the representation of code in the brain: an fMRI study of code review and expertise. In: *2017 IEEE/ACM 39th International Conference on Software Engineering (ICSE).* IEEE. p. 175–186.
- Friedrich R, Friederici AD. 2009. Mathematical logic in the human brain: syntax. *PLoS One.* 4:e5599.
- Goulas A, Uylings HB, Stiers P. 2012. Unravelling the intrinsic functional organization of the human lateral frontal cortex: a parcellation scheme based on resting state fMRI. *J Neurosci.* 32:10238–10252.
- Grover S, Cooper S, Pea R. (2014). Assessing computational learning in K-12. *Paper presented at the proceedings of the 2014 conference on Innovation & Technology in Computer Science Education.*
- Hao X, Wang X, Song Y, Kong X, Liu J. 2018. Dual roles of the hippocampus and intraparietal sulcus in network integration and segregation support scene recognition. *Brain Struct Funct.* 223:1473–1485.
- Hobeika L, Diard-Detoeuf C, Garcin B, Levy R, Volle E. 2016. General and specialized brain correlates for analogical reasoning: a meta-analysis of functional imaging studies. *Hum Brain Mapp.* 37:1953–1969.
- Humphreys GF, Lambon Ralph MA. 2015. Fusion and fission of cognitive functions in the human parietal cortex. *Cereb Cortex.* 25:3547–3560.
- Kalelioglu F, Gulbahar Y, Kukul V. 2016. A framework for computational thinking based on a systematic research review. *Balt J Modern Comput.* 4:583.
- Kazakoff ER, Bers MU. 2014. Put your robot in, put your robot out: sequencing through programming robots in early childhood. *J Educ Comput Res.* 50:553–573.
- Kleinmintz OM, Ivancovsky T, Shamay-Tsoory SG. 2019. The two-fold model of creativity: the neural underpinnings of the generation and evaluation of creative ideas. *Curr Opin Behav Sci.* 27:131–138.
- Klimesch W. 1999. EEG alpha and theta oscillations reflect cognitive and memory performance: a review and analysis. *Brain Res Rev.* 29:169–195.
- Lange T, Roth V, Braun ML, Buhmann JM. 2004. Stability-based validation of clustering solutions. *Neural Comput.* 16:1299–1323.
- Latini-Corazzini L, Nesa MP, Ceccaldi M, Guedj E, Thinus-Blanc C, Cauda F, Dagata F, Péruch P. 2010. Route and survey processing of topographical memory during navigation. *Psychol Res.* 74:545–559.
- Li Y, Kong F, Ji M, Luo Y, Lan J, You X. 2019. Shared and distinct neural bases of large- and small-scale spatial ability: a coordinate-based activation likelihood estimation meta-analysis. *Front Neurosci.* 12:1–20.
- Mannila L, Dagiene V, Demo B, Grgurina N, Mirolo C, Rolandsson L, Settle A. 2014. Computational thinking in K-9 education. *Proceedings of the Working Group Reports of the 2014 on Innovation & Technology in Computer Science Education Conference.* p. 1–29.
- McLaren DG, Ries ML, Xu G, Johnson SC. 2012. A generalized form of context-dependent psychophysiological interactions

- (gPPI): a comparison to standard approaches. *Neuroimage*. 61:1277–1286.
- Milivojevic B, Johnson BW, Hamm JP, Corballis MC. 2003. Non-identical neural mechanisms for two types of mental transformation: event-related potentials during mental rotation and mental paper folding. *Neuropsychologia*. 41:1345–1356.
- Owen AM, McMillan KM, Laird AR, Bullmore ET. 2005. N-back working memory paradigm: a meta-analysis of normative functional neuroimaging. *Hum Brain Mapp*. 25:46–59.
- Papert S. 1990. *Children, computers and powerful ideas*. New York: Basic Books.
- Peppler KA, Warschauer M. 2011. Uncovering literacies, disrupting stereotypes: examining the (dis) abilities of a child learning to computer program and read. *Int J Learn Media*. 3:15–41.
- Peterson BS, Skudlarski P, Gatenby JC, Zhang HP, Anderson AW, Gore JC. 1999. An fMRI study of Stroop word-color interference: evidence for cingulate subregions subserving multiple distributed attentional systems. *Biol Psychiatry*. 45:1237–1258.
- Prabhakaran V, Smith JAL, Desmond JE, Glover GH, Gabrieli JDE. 1997. Neural substrates of fluid reasoning: an fMRI study of neocortical activation during performance of the Raven's progressive matrices test. *Cogn Psychol*. 33:43–63.
- Raghavachari S, Kahana MJ, Rizzuto DS, Caplan JB, Kirschen MP, Bourgeois B, Madsen JR, Lisman JE. 2001. Gating of human theta oscillations by a working memory task. *J Neurosci*. 21:3175–3183.
- Roman-Gonzalez M, Perez-Gonzalez J-C, Jimenez-Fernandez C. 2017. Which cognitive abilities underlie computational thinking? Criterion validity of the computational thinking test. *Comput Hum Behav*. 72:678–691.
- Rottschy C, Langner R, Dogan I, Reetz K, Laird AR, Schulz JB, Fox PT, Eickhoff SB. 2012. Modelling neural correlates of working memory: a coordinate-based meta-analysis. *Neuroimage*. 60:830–846.
- Selby C, Woollard J. 2013. Computational thinking: the developing definition. *The 18th Annual Conference on Innovation and Technology in Computer Science Education, Canterbury*.
- Sharp D, Bonnelle V, De Boissezon X, Beckmann C, James S, Patel M, Mehta MA. 2010. Distinct frontal systems for response inhibition, attentional capture, and error processing. *Proc Natl Acad Sci*. 107:6106–6111.
- Shulman GL, Fiez JA, Corbetta M, Buckner RL, Miezin FM, Raichle ME, Petersen SE. 1997. Common blood flow changes across visual tasks decreases in cerebral cortex. *J Cogn Neurosci*. 9:648–663.
- Siegmund J, Kaestner C, Apel S, Parnin C, Bethmann A, Leich T, Saake G, Brechmann A. 2014. Understanding understanding source code with functional magnetic resonance imaging. *Proceedings of the 36th ACM/IEEE International Conference on Software Engineering*. 378–389.
- Taylor JC, Wiggett AJ, Downing PE. 2007. Functional MRI analysis of body and body part representations in the extrastriate and fusiform body areas. *J Neurophysiol*. 98:1626–1633.
- Tedre M, Denning PJ editors. *The long quest for computational thinking*. *Proceedings of the 16th Koli Calling International Conference on Computing Education Research*; 2016. p. 120–129.
- Tesche CD, Karhu J. 2000. Theta oscillations index human hippocampal activation during a working memory task. *Proc Natl Acad Sci USA*. 97:919–924.
- Vilberg KL, Rugg MD. 2008. Memory retrieval and the parietal cortex: a review of evidence from a dual-process perspective. *Neuropsychologia*. 46:1787–1799.
- Vincent JL, Kahn I, Snyder AZ, Raichle ME, Buckner RL. 2008. Evidence for a frontoparietal control system revealed by intrinsic functional connectivity. *J Neurophysiol*. 100:3328–3342.
- Wagner AD, Shannon BJ, Kahn I, Buckner RL. 2005. Parietal lobe contributions to episodic memory retrieval. *Trends Cogn Sci*. 9:445–453.
- Wang X, Zhen Z, Song Y, Huang L, Kong X, Liu J. 2016. The hierarchical structure of the face network revealed by its functional connectivity pattern. *J Neurosci*. 36:890–900.
- Weintrop D, Beheshti E, Horn M, Orton K, Jona K, Trouille L, Wilensky U. 2016. Defining computational thinking for mathematics and science classrooms. *J Sci Educ Technol*. 25:127–147.
- Whitfield-Gabrieli S, Nieto-Castanon A. 2012. Conn: a functional connectivity toolbox for correlated and anticorrelated brain networks. *Brain Connect*. 2:125–141.
- Wing JM. 2006. Computational thinking. *Communications of the Acm*. 49:33–35.
- Wing JM. 2008. Computational thinking and thinking about computing. *IEEE International Symposium on Parallel & Distributed Processing*. 2008. p. 28–28.
- Winlove CI, Milton F, Ranson J, Fulford J, MacKisack M, Macpherson F, Zeman A. 2018. The neural correlates of visual imagery: a co-ordinate-based meta-analysis. *Cortex*. 105:4–25.
- Xu S, Humphreys GW, Heinke D. 2015. Implied Actions Between Paired Objects Lead to Affordance Selection by Inhibition. *Journal of Experimental Psychology: Human Perception and Performance*. 41:1021–1036.
- Yan C-G, Wang X-D, Zuo X-N, Zang Y-F. 2016. DPABI: data processing & analysis for (resting-state) brain imaging. *Neuroinformatics*. 14:339–351.
- Yeh MKC, Gopstein D, Yan Y, Zhuang Y, Ieee. 2017. Detecting and comparing brain activity in short program comprehension using EEG. *2017 IEEE Frontiers in Education Conference*.
- Yeo BTT, Krienen FM, Sepulcre J, Sabuncu MR, Lashkari D, Hollinshead M, Roffman JL, Smoller JW, Zoller L, Polimeni JR, et al. 2011. The organization of the human cerebral cortex estimated by intrinsic functional connectivity. *J Neurophysiol*. 106:1125–1165.
- Yu M, Wu Z, Luan M, Wang X, Song Y, Liu J. 2018. Neural correlates of semantic and phonological processing revealed by functional connectivity patterns in the language network. *Neuropsychologia*. 121:47–57.
- Zhou X, Li M, Li L, Zhang Y, Cui J, Liu J, Chen C. 2018. The semantic system is involved in mathematical problem solving. *Neuroimage*. 166:360–370.