the british
psychological society
promoting excellence in psychology

Journal of Neuropsychology (2020)
© 2020 The British Psychological Society
www.wileyonlinelibrary.com

Left posterior prefrontal regions support domain-general executive processes needed for both reading and math

Kai Wang[1,2]*, Marie T. Banich[2,3]*, Andrew E. Reineberg[3,4], Daniel R. Leopold[3], Erik G. Willcutt[3,4], Laurie E. Cutting[5], Stephanie N. Del Tufo[6], Lee A. Thompson[7], John Opfer[8], Frank J. Kanayet[8], Zhong-Lin Lu[9,10,11,12] and Stephen A. Petrill[8]

[1]School of Psychology, South China Normal University, Guangzhou, Guangdong, China
[2]Institute of Cognitive Science, University of Colorado Boulder, Boulder, Colorado, USA
[3]Department of Psychology and Neuroscience, University of Colorado Boulder, Boulder, Colorado, USA
[4]Institute for Behavioral Genetics, University of Colorado Boulder, Boulder, Colorado, USA
[5]Peabody College of Education and Human Development, Vanderbilt University, Nashville, Tennessee, USA
[6]School of Education, University of Delaware, Newark, Delaware, USA
[7]Department of Psychological Sciences, Case Western Reserve University, Cleveland, Ohio, USA
[8]Department of Psychology, The Ohio State University, Columbus, Ohio, USA
[9]Division of Arts and Sciences, NYU Shanghai, Shanghai, China
[10]NYU-ECNU Institute of Brain and Cognitive Neuroscience at NYU Shanghai, Shanghai, China
[11]Center for Neural Science, New York University, New York, New York, USA
[12]Department of Psychology, New York University, New York, New York, USA

Substantial evidence has suggested that reading and math are supported by executive processes (EP). However, to date little is known about which portion of the neural system underpinning domain-general executive skills works to support reading and math. In this study, we aimed to answer this question using fMRI via two complementary approaches. First, imaging data were acquired whilst a sample of 231 adolescents performed each of three separate tasks designed to assess reading comprehension, numerical magnitude

---

*Correspondence should be addressed to Kai Wang, Room 711, School of Psychology, South China Normal University, Guangzhou 510631, China. (email: kaiwang@m.scnu.edu.cn) and Marie T. Banich, Institute of Cognitive Science and Department of Psychology & Neuroscience, E230 Muenzinger Hall, University of Colorado Boulder, Boulder, Colorado 80309-0345, USA. (email: Marie.Banich@colorado.edu).

DOI:10.1111/jnp.12201

estimation, and EP in working memory (WM), respectively. With careful task designs and conjunction analyses, we were able to isolate cross-domain brain activity specifically related to EP, as opposed to lower-level domain-general processes (e.g., visual processing). Second, the meta-analytic tool Neurosynth was used to independently identify brain regions involved reading, math, and EP. Using a combination of forward and reverse statistical inference and conjunction analyses, we again isolated brain regions specifically supporting domain-general EP. Results from both approaches yielded overlapping activation for reading, math, and EP in the left ventrolateral prefrontal cortex, left inferior frontal junction, and left precentral gyrus. This pattern suggests that posterior regions of the prefrontal cortex, rather than more central regions such as mid-DLPFC, play a leading role in supporting domain-general EP utilized by both reading and math.

Complex goal-oriented behaviours rely on executive processes (EP) that purposely exert top–down control over basic functions like motor and perceptual processes (Banich, 2009; Diamond, 2013; Miller & Cohen, 2001). EP operate across various task contexts and are thought to be domain-general (Duncan, 2010; Rajah, Ames, & D'Esposito, 2008). For this reason, there has been an enduring interest in the role of EP in primary academic skills like reading and math. To date, substantial evidence has shown that both reading and math are supported by certain aspects of EP. Individual differences in EP predict both current (Best, Miller, & Naglieri, 2011; Jacob & Parkinson, 2015; Lan, Legare, Ponitz, Li, & Morrison, 2011; Yeniad, Malda, Mesman, van Ijzendoorn, & Pieper, 2013) and future (Bull, Espy, & Wiebe, 2008; Jacob & Parkinson, 2015) levels of reading and math performance. Moreover, deficits in EP are related to low achievement in both academic domains (Biederman *et al.*, 2004; Rose, Feldman, & Jankowski, 2011). Whilst these behavioural studies provide key evidence of the cognitive overlap between reading and math, little, however, is known about the neural mechanisms underlying these domain-general EP that support both reading and math skills.

The fronto-parietal network is known to support EP (Cole, Repovš, & Anticevic, 2014; Vincent, Kahn, Snyder, Raichle, & Buckner, 2008). It is unclear, however, exactly *which* portions of this extended network underlie processes required for both reading and math. Whilst one always has to be cautious about making reverse inferences, identifying these brain regions would provide insights into the likely EP that are most likely jointly associated with reading and math. As such, this information could enhance current theoretical models and interpretations of behavioural findings that consistently show certain EP – most often those associated working memory (WM) – are also associated with reading and math ability.

Whilst the conceptualization of executive processes as used in studies relevant to reading and reading abilities is quite varied (Church *et al.*, 2019) and similar issues hold with regard to mathematical abilities (Bull & Lee, 2014), there are at least two aspects of executive function that are likely to influence both reading and math. A prominent model of executive function, the unity and diversity model (Friedman & Miyake, 2017; Miyake *et al.*, 2000) proposes that there is a unitary factor, referred to as common executive function (EF), underlying all executive function tasks, much as *g* is thought to undergird performance across intelligence tasks. This common EF factor is posited to represent the ability to actively hold task goals in mind, especially in the face of distraction. Performance on tasks of inhibitory control loads completely on that factor, as they often involve maintaining a task set in the face of distraction or pre-potent response tendencies. Once the common EF factor is accounted for, there are at least two distinct executive function subfactors, reflecting the diversity aspect of the model. One is a working memory updating-specific factor that is linked to the ability to update the contents of working memory. The other is a task-switching specific factor that provides the ability to switch between task sets.

Both common EF and the working memory updating-specific factor can be reasonably thought to influence reading and math. Cirino *et al.* (2019) found that reading comprehension in middle childhood is influenced both by common EF and working memory updating. Whilst we know of no study that has taken a similar approach in distinguishing between these two components of EF (i.e., common EF and working memory updating-specific EF) in relation to math, the empirical work that does exist suggests a similar influence of these two factors. For example, Cragg, Keeble, Richardson, Roome, and Gilmore (2017) found that tasks that measure the updating of working memory predict mathematical achievement in children, youth, and young adults, and that these tasks along with inhibitory tasks (which load on common EF) predict factual knowledge and procedural skill in math.

At present, the exact neural structures associated with the factors in the unity and diversity model remain unclear. However, there is consensus across the field that lateral prefrontal cortex (LPFC) plays a central role in executive processes by biasing information processing towards task-related representations or processes (Banich, 2009; Miller & Cohen, 2001). This top–down biasing process is assumed to be able to affect information processing in multiple lower-level sensory or motor modalities and relies on connectivity of dorsolateral prefrontal regions to target regions (Depue, Orr, Smolker, Naaz, & Banich, 2015).

Additional neuroimaging studies provided finer-grain predictions on candidate regions of domain-general EP. The mid-DLPFC is implicated in the manipulation and monitoring of information in WM, especially to help resolve interference from task-irrelevant information (Barbey, Koenigs, & Grafman, 2013; Petrides, 2000). In contrast, the inferior frontal junction (IFJ; Muhle-Karbe *et al.*, 2016) is shown to be involved in the identification of information relevant to task goals and demands (Muhle-Karbe *et al.*, 2016). Additionally, premotor regions are implicated in the articulatory aspects of WM (Price, 2012), whilst ventrolateral prefrontal cortex (VLPFC) has been linked to accessing and selecting relevant information from memory (Badre & Wagner, 2007), especially in regard to selecting amongst relevant potential words (Snyder *et al.*, 2010). Finally, the prefrontal cortex is known to have some degree of functional lateralization. For example, right inferior frontal cortex is more involved in inhibitory control (Aron, Robbins, & Poldrack, 2004) whilst left inferior frontal cortex plays a more prominent role in processing of linguistic information (Price, 2012). Which of these areas shows activations across all three tasks, therefore, is likely to provide insights into the types of EP that are jointly required by reading and math.

In fact, there has been some evidence supporting the idea that numerous regions involved in EP are also active during tasks that tax reading and/or math mechanisms. With regard to reading, activation is observed in multiple regions of the left PFC known to be involved in executive processes (Fedorenko & Thompson-Schill, 2014; Ferstl, Neumann, Bogler, & von Cramon, 2008). Moreover, studies in which both EP and reading tasks were administered found overlapping activations in the middle and inferior PFC as well as anterior cingulate cortex (January, Trueswell, & Thompson-Schill, 2009; Ye & Zhou, 2009). With regard to math, neuronal activity in the PFC and parietal lobe is thought to be involved in the encoding and processing of numerical magnitude (Nieder, 2016). More interestingly, with a factorial design, Ansari, Fugelsang, Dhital, and Venkatraman (2006) reported an interaction of the EP and number estimation processes in the left frontal lobe, which was shown recently in another study to be predictive of levels of math achievement (Wilkey & Price, 2018). These lines of evidence suggest that math and reading may potentially rely on shared neural mechanisms involved in EP.

In this study, we chose to examine EP mechanisms in a sample of late adolescents, an understudied age with regard to these abilities. At this critical developmental stage, individual's basic cognitive abilities including basic reading and math ability are relatively well-established (Best & Miller, 2010; Chall, 1983b; Halberda, Ly, Wilmer, Naiman, & Germine, 2012). In contrast, EP and the prefrontal regions that support them are relatively less mature (Casey, Jones, & Hare, 2008; Gogtay *et al.*, 2004), and their development involves both changes in the specific regions of prefrontal cortex engaged and their degree of activation (Andrews-Hanna *et al.*, 2011). Hence, from a developmental perspective, the exact region of prefrontal cortex that may be engaged in EP along with reading and math is difficult to predict.

The present investigation used two complementary sources of data to examine the EP regions jointly employed by math and reading. In an empirical approach, participants were imaged whilst performing each of three different tasks: a sentence-level reading comprehension task, a numerical magnitude estimation task, and a task requiring EP that act on WM. Conjunction analyses in this within-subject design were then used to determine which brain regions supporting domain-general EP also contribute to reading and math performance. Due to the time constraints of the overall neuroimaging study, however, we could only assess each of the constructs of interest (i.e., EP, reading, and math) using one specific and limited task.

To provide insights as to the generalizability of these results, the second complementary approach used Neurosynth, a powerful meta-analytic tool designed to interrogate patterns of brain activation associated with particular psychological functions (http://www.neurosynth.org/; Yarkoni, Poldrack, Nichols, Van Essen, & Wager, 2011). In this approach, patterns of brain activation are deduced from hundreds of studies that more broadly assess aspects of mathematical processing, reading, and EP in WM, respectively. Not only can this approach potentially provide converging evidence for the results from the empirical approach, but it can also capture a broader array of related functions across each of these three domains than is possible within a single neuroimaging paradigm. We then present the regions identified by both approaches.

Another impediment to identifying EP underlying reading and math is that they are not the only domain-general processes potentially shared between reading and math. Other domain-general processes may include the decoding of visual symbols, likely relying upon portions of ventral occipital-temporal cortex (For a review, see Hannagan, Amedi, Cohen, Dehaene-Lambertz, & Dehaene, 2015), the linkage of such perceptual visual information to an abstract representation, which relies on inferior parietal regions (Seghier, 2013), and linkage of those abstract representations to an amodal semantic system, thought to rely partly on anterior temporal regions (Simmons & Martin, 2009). Hence, the methods we used were designed to intentionally separate the contributions of these lower-level domain-general processes shared between reading and math from the EP shared between reading and math (refer to Figure 1).

Determining which EP-related brain regions might support both reading and math is a non-trivial task, as there are challenges associated with such an undertaking. First, ideally all three domains of cognition (i.e., EP, reading, and math) should be included in the same scanning session to allow for a controlled within-subjects test. Second, study designs must isolate the higher-level EP from the confounding lower-level domain-general processes (see Figure 1 and Methods). Third, a large sample is required to obtain stable results.

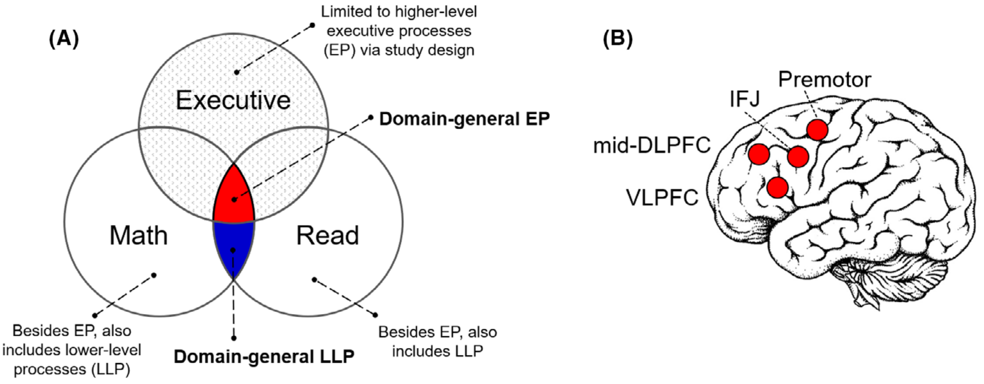


Figure 1. Logic of the analytic approach to isolate the neural underpinnings of the domain-general executive processes jointly employed by reading and math skills. (A) In both empirical and meta-analytic investigations, by *excluding lower-level processes (LLP) in the executive map (see Method Section 1), the conjunction of all three maps can capture the domain-general executive processes.* (B) Potential prefrontal regions supporting executive processes for reading and math include middle dorsolateral (mid-DLPFC) linked to manipulating and monitoring the contents of working memory, ventrolateral (VLPFC) prefrontal cortex linked to accessing and selection of information from long-term memory, premotor regions involved in articulatory functions, and the inferior frontal junction (IFJ) involved in selection of task-relevant information.

Likely due to these difficulties, we know of no imaging study that has determined those aspects of EP common to reading and math.

In sum, in this study we adopted two complimentary approaches to examine brain regions that underlie EP supporting both reading and math skills in a domain-general fashion (Figure 1). Based on previous studies, we expect that one or multiple areas within the left prefrontal lobe (middle, ventral, posterior, and premotor PFC) will be identified across both approaches, although the exact area is difficult to predict *a priori*.

Method

Overall approach

The overall approach used with our empirical data and with Neurosynth was similar. The goal was to isolate brain regions activated by both reading comprehension and math functions (a two-way conjunction) and then determine which of these is also activated by EP in WM (a three-way conjunction). Critically, the maps used for reading and math were chosen to identify all regions that contribute to both these cognitive domains, including regions involved in lower-level domain-general processing (e.g., visual form processing) as well as those involved in higher-level EP (Figure 1). In contrast, the map used for EP in WM was designed to isolate activity specific to higher-level EP excluding lower-level processes. The three-way conjunction then identifies the subset of the two-way regions (between reading and math) that are specifically also involved in higher-level EP (red regions in Figure 1). In contrast, non-EP activations related to both reading and math can then be obtained by excluding the three-way regions from the results of the two-way conjunction analyses (blue regions in Figure 1). To achieve these goals, the reading and math maps were derived in a manner distinct from those for EP (see details below).

We used the N-back task as a measure of executive processes that must act on WM, because it requires not only the encoding and temporal storage of information in WM, but a variety of executive processes including the continuous updating of information in WM and the inhibition of irrelevant stimuli and responses (Rac-Lubashevsky & Kessler, 2016). Hence, the executive manipulation of information in WM seems to be critical for this task. Supporting this idea, it has been shown that N-back task correlates only weakly with WM capacity, meanwhile, its correlations with tasks or abilities involving EP are stronger (Gajewski, Hanisch, Falkenstein, Thönes, & Wascher, 2018; Jaeggi, Buschkuehl, Perrig, & Meier, 2010; Kane, Conway, Miura, & Colflesh, 2007). Also suggesting that the N-back task taps a variety of executive processes, meta-analyses indicate that it engages a wide range of frontal regions including left middle frontal and left inferior frontal gyrus, which are commonly activated across a wide variety of executive function tasks (Mencarelli *et al.*, 2019; Wang, He, Wu, Zhang, Jin & Li, 2019).

A naturalistic reading comprehension task was used to assess participants' reading processes. A numerical line estimation task was used for the math domain (Siegler & Opfer, 2003). This task is able to test numerical processes independent of world knowledge and has been shown to be correlated with math performance (Schneider *et al.*, 2017). To assess executive processes linked to working memory, participants were given an N-back task. The results of this design can thus provide insight into the specific cognitive and neurocognitive mechanisms of domain-general EP that contribute to reading and math.

Empirical investigation

Participants

Participants for the study were drawn from the Learning Disabilities Innovation Hub (LDIH) sample, which is based on a collaboration between The Ohio State University, Case Western Reserve University, the University of Colorado Boulder, and Vanderbilt University. Participants for the study consisted of 231 adolescents (mean age = 16.83, *SD* = 1.49, 55.8% female, age range 13.92–20.75, 114 MZ twins, 102 same-sex DZ twins and 15 singletons) recruited into the LDIH from the Western Reserve Reading and Math Project (Soden *et al.*, 2015) and the Colorado portion of the International Longitudinal Twin Study of Early Reading Development (Christopher *et al.*, 2015). Such a sample size should provide more than adequate power, as this sample size is at least 5–10 times the typical sample size for neuroimaging studies. Data collection was accomplished at three universities: The Ohio State University, Case Western Reserve University and the University of Colorado Boulder. This sample is unselected for reading or math disabilities. The influence of participants' non-independence (due to the inclusion of twins) on statistical inference was controlled for by defining exchangeability blocks in permutation tests (see section 'Higher-level GLM analyses'). Adolescents were chosen for this investigation as their reading and math abilities are relatively mature (Chall, 1983b; Halberda *et al.*, 2012). All participants had normal or corrected-to-normal vision and had no history of neurological disorders. Informed consent was obtained from each individual (for those older than 18) or from that individual's parent or guardian (and his or her assent for individuals younger than 19) prior to participation. Informed consent was obtained from each family prior to participation, and all study procedures were approved by the Institutional Review Board at The Ohio State University, the Case Western Reserve University Social/Behavioral Institutional Review Board and University of Colorado Boulder Institutional Review Board.

Parents of twin pairs were contacted by The Ohio State University or by University of Colorado Boulder by their preferred mode of contact as indicated during their prior research participation. Aims of the study were described, and families who responded were provided with a description of the project's overview, purpose, and data collection procedures. 80% of twins previously recruited at The Ohio State University were initially nominated by schools in Greater Cleveland, Columbus, and Cincinnati metropolitan areas, as well as throughout Ohio and Western Pennsylvania. Schools sent information packets home with parents of twins who had entered kindergarten but not yet completed first grade. Additional families were recruited via Ohio State birth records, mother of twin clubs, and media advertisement. Twins previously recruited by the University of Colorado Boulder were initially ascertained through the Colorado Twin Registry, which contains records for all twin births in Colorado besides families that explicitly asked to not be contacted. All twins were said to be in their final year of pre-school at the time of recruitment, and 86% of parents approached by telephone agreed to participate.

The racial and ethnic composition of the sample is 86.7% White, 2.9% African American, 2.9% Hispanic, 0.6% Asian, and 2.3% Native American, Pacific Islander, or Multiple Races, with 4.6% not reporting. Mean parental education was 15.67 years (*SD* = 2.05 years; range = 10–20 years).

Procedure

Functional MRI was utilized to assess patterns of brain activation whilst participants performed a task from each of three domains of interest: EP in WM (via an N-back task), reading (via a sentence-level reading comprehension task), and math (via a number line magnitude estimation task).

For the tasks assessing reading and math, contrasts were chosen that would enable capture of all relevant brain activation whether or not it is related to EP. To that end, the main contrast of interest was between an active task condition (passage reading for reading or mapping a number's magnitude onto an axis for math) and a passive control condition (symbol reading or fixation on an axis, respectively). By using a relatively basic (i.e., low-level) control condition, the full range of brain regions engaged in each of these tasks, from lower-level sensory processing to higher-order processes, could be identified. Differently, however, for the N-back task, a contrast between two active task conditions (2-back vs. 0-back) with identical perceptual stimuli was used. This contrast allows the specific identification of regions involved in EP without inclusion of regions involved in lower-level non-EP domain-general processes (see Figure 2).

Image acquisition. All data were collected on 3 Tesla Siemens TIM Trio MRI scanners with a 12-channel head coil at three imaging centres. Whole-brain structural images were collected with a high-resolution 3D magnetization prepared rapid gradient echo (MPRAGE) sequence (TR = 3 ms, TI = 755 ms, Flip Angle = 8°; Slice thickness = 1.2 mm; FOV = 26 cm; Matrix size: 256 × 256.). Functional BOLD (blood oxygenation level dependent) images were collected using gradient echo T2*-weighted echo planar imaging (EPI); (axial acquisition geometry for the NLE and reading task, SENSE factor = 2.5, TR = 2.2 s, TE = 25 ms, flip angle = 72°, 42 slices, 3 mm slice thickness, 1mm gap, bandwidth = 2,604 Hz/pixel; for the N-back task, the TE was 28 ms). The slices were oriented obliquely along the AC–PC line. The first four volumes from each functional run were discarded to allow for magnetic field equilibration. DTI data were also

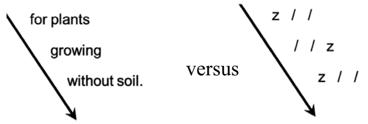
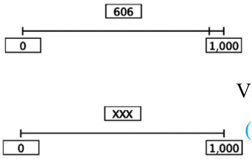
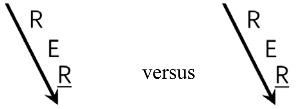
Design	Empirical investigation		Meta-analytic	
	Contrast	Processes involved	Statistic inference type	Processes involved
Domains reading	 (Paragraph reading) (Symbol reading)	i. Reading Comprehension ii. LLP iii. EP	Forward	Any process associated with reading
Math	 (Number line estimation) Vs. (Fixation on axis)	i. Numerical Estimation ii. LLP iii. EP	Forward	Any process associated with math
Executive	 (2-back task) (0-back task)	EP (as stimuli are the same in both conditions)	Reverse	EP specifically

Figure 2. Contrasts (empirical data) and statistical inference maps (Neurosynth) used in conjunction analyses to examine the overlap of brain areas involved in reading, math, and executive processes. By excluding lower-level domain-general processes in the executive map (N-back task), the conjunction of all three maps could capture the domain-general aspect of executive processes. EP = executive processes LLP = lower-level processes. (Left-hand side) Empirical Investigation. Each participant performed three tasks. Top Row: Reading task. Activation during paragraph reading was compared to reading symbols. Middle Row: Participants determined whether the hash mark on a number line matched a displayed numerical value vs. fixation on the axis. Bottom Row: Activation whilst detecting whether a letter was identical to the letter presented two items back was compared to responding to underlined letters. (Right-hand side) Meta-analytic Investigation. For both the reading and math topics, forward inference maps in Neurosynth (see methods) were used to identify all brain regions significantly activated in studies associated with the topics of reading and math, respectively. For the EP topic, a reverse inference was used to detect only those brain regions that are significantly associated with the particular topic above and beyond other topics.

collected but are not discussed here. The structural image was acquired first, followed by the functional scans for the reading task and the number-line estimation task, respectively, with a DTI scan next and then the N-back task. Each task is described in more detail below. All participants performed the tasks in the same order as one goal of the larger project, described in a separate report (Wang *et al.*, 2019), was to examine individual differences in pattern of brain activation across our sample with regard to reading and executive abilities.

Task paradigms

For all tasks, stimuli were projected using an active video projector and presented on a screen at the foot of the scanner bed, viewed through a mirror placed above the participant's head. A fibre-optic response box with two buttons was used for responses. In addition, at the conclusion of these three tasks individuals performed a 6-min resting-state scan.

Reading task. This task utilized a block design with three distinct conditions. For all three conditions, however, the participant's task was to press a button to indicate when two consecutive screens contained repeated content. Such repetition occurred randomly.

In the passage reading condition, the active condition of interest, an initial blank screen lasting 1,000 ms, was followed by a series of 1–5 word phrases that formed a meaningful paragraph. Each phrase was presented on a separate trial. As done in Aboud, Bailey, Petrill, and Cutting (2016), we allowed 550 ms for each content word (i.e., nouns, verbs, adjectives, and adverbs that relay semantic information) and 275 ms for each function word (i.e., pronouns, conjunctions, prepositions, articles, and auxiliary verbs that relay information about syntactic relationships; see Diaz *et al.*, 2009). For example, if one screen contained one content word and one function word, it would be displayed for 825 ms. In the symbol block, the baseline task against the paragraph condition was compared, a series of symbol/letter triads (e.g., 'z/ /', '/ z/' or '/ / z') were presented. In both conditions, participants pressed a button if subsequent screens contained the same items. There were 21–41 trials in a reading block and 14–16 trials in a symbol block. The task also contained a word condition, which is not of interest here. In this condition, the stimuli display and the task were the same as the meaningful paragraph condition, except just a series of content words was displayed.

There were three runs in all. The sequence of the first two runs was 'Paragraph-symbol-paragraph-symbol-word-symbol', and the sequence of the last run was 'Word-symbol-paragraph-symbol-paragraph-symbol'. The two paragraphs within a run belonged to a single passage. The first paragraph served to introduce the topic whilst the second elaborated on a particular detail of the subject matter. The passages were all expository and included the following topics: Hang Gliding, Wrasses, and Hydroponics.

To ensure that individuals were comprehending the paragraphs, immediately after the second paragraph block, participants were presented with two picture-judgement trials. In these trials, a picture was presented for eight seconds followed by a response screen lasting for four seconds during which participants determined whether a picture matched the content of the two paragraphs by making a left or right key-press. The contrast of interest for this task was the paragraph condition minus the symbol condition.

Number-line estimation (NLE) Task – This task utilized an event-related design that involved two conditions – a number-line estimation (NE) condition and a luminance estimation (LE) condition. The two active conditions were intermixed in an event-related manner across two runs of 72 trials each. However, just activation for the number-line estimation task was examined in the present study.

Each trial began with a cue informing the participants of the trial condition that lasted for a jittered time period of 1,500–9,200 ms. The analyses of the fMRI data contrast activation for the NE trials (described below) against the baseline provided by this cue. During the cue period, a black horizontal line was presented on a white screen to visually map out a spatial reference frame from left to right. For NE trials, two numerals were presented below the horizontal line, 0 at the left end, and 1,000 on the right. For LE trials, the numerals were replaced with square boxes, with the box at the left end filled white, and the box at the right end filled black. Above the number line, there was a fixation box with 'XXX' printed inside. The fixation box appeared ¾ of the way from the bottom of the screen.

During the NE trial, after the jittered cue period, a probe appeared inside the fixation box for a period of 1,000 ms. The probe could be one of six different Arabic numerals (5, 78, 150, 606, 725, or 938). After the probe period, a hatch mark was displayed on the horizontal line for 2,000 ms. The hatch mark could appear at one of three different positions with equal probability. The distance between the left end and the hatch mark could correspond to: (1) the linear value of the probe, (2) the logarithmic value of the probe, or (3) the mid-point between the linear and logarithmic values. Participants were required to make a yes/no decision as to whether the position of the hatch mark matched the numeral, and they were required to answer before the end of the 2,000 ms in which the hatch mark appeared on screen.

During LE trials, which were not of interest for the present purposes, instead of using numerals, the probes were replaced with luminance patches in greyscale. On LE trials, the probes were constructed as square boxes with luminance levels that corresponded to the values of the numerals in the NE condition. Shades of grey were constructed using the RGB colour system. Each luminance patch was created by selecting an RGB value in the range from 0 to 255 that corresponded to each of the 6 numerals tested in the NE trials. This was accomplished using the following formula: RGB luminance = (1,000 – numeral) * (255/ 1,000)).

N-back task – This task consisted of 2-back and 0-back conditions that were presented in a blocked periodic design alternating between the two conditions. The task session consisted of one run of ten blocks. At the beginning of each block, a text cue was presented for 2.2 s or 4.4 s (25% of the time) to indicate the condition of the following block. In each block, a series of 14 letters in white font was visually presented on a black screen for 0.75 s each followed by a blank screen for 1.45 s. Half of the letters (randomly selected) were underlined. In the control condition (0-back condition), participants pressed a key whenever an underlined letter appeared on the current trial. In the 2-back condition, participants pressed a key if the current letter matched the one presented two trials ago.

Imaging analysis

Pre-processing. Image processing and data analysis were implemented using the FSL package (analysis group, FMRIB, Oxford, UK, http://www.fmrib.ox.ac.uk/fsl/). Standard pre-processing was applied: MCFLIRT, linear slice time correction/motion correction, BET, brain extraction, time-series pre-whitening, high pass filter (0.01 Hz), and registration and spatial normalization to the Montreal Neurological Institute (MNI) 152-T1 2-mm template. The individual's functional images were first registered to their high-resolution MPRAGE scans via a 6-parameter linear registration, and the MPRAGE images were in turn registered to the MNI template via a 12-parameter linear registration. These registrations were combined in order to align the functional images to the template. Functional images were resampled into standard space with 2-mm isotropic voxels and were smoothed with a Gaussian kernel of an 8-mm full width at half-maximum.

FMRI time-series data were analysed using the general linear model (GLM) after filtering low-frequency noise, correcting for temporal autocorrelation using an autoregressive AR(1) model, and convolving the stimulus function with a double-gamma hemodynamic response function. Within the GLM approach, six linear head movement parameters (*X*, *Y*, *Z*, roll, pitch, and yaw) and their squared value were included as confound regressors. Task-specific regressors at the lower-level GLM are described below.

Quality assurance (QA) was performed before any further analysis of fMRI data. This QA involved examining the data in a number of distinct manners. All data were examined for their brain coverage. Data that failed to cover the entire cerebral cortex were excluded from further analysis. However, we allowed for some coverage loss in the inferior part of cerebellum. Second, data with a rotation <2 degrees and a translation <2 mm within a scanning session were directly accepted for subsequent analyses. Data failing to meet these rotation and translation standards were further evaluated by two of the authors independently through three methods. First, we examined the pattern of the movement plot generated by MCFLIRT of FSL (Jenkinson, Bannister, Brady, & Smith, 2002). Plots that showed frequent and sharp peaks of movement were considered worse than those showed continued but slow movement. Second, we examined the severity of movement-related effects (Poldrack, Mumford, & Nichols, 2011) on the statistic map of interested contrast (N-back task, 2-back vs. 0-back; NLE, number size parametric effect; reading task, passage vs. symbol reading). Besides the standard GLM described above, for all fMRI data failing to meet our rotation and translation standards, two more GLMs were run with two additional types of confound regressors respectively, one using the metrics of frame displacement (fd) and the other using the metrics of RMS intensity difference of volume *N* to volume *N* + 1 (dvars). The FSL tool of fsl_motion_outliers was used to generate both type of regressors. In order to check the effect of these additional regressors, we examined the severity of movement-related effects for both results of the standard GLM and results of the GLMs with additional confound regressors, Third, we visually assessed the quantity and severity of images distortion due to head movement by examining the all images within a scanning session played as a movie using fslview. Each of these criteria was graded on a scale of 1–4 (1 = good, 4 = bad). Any data had an average score larger than 3 by either author were not included in further analysis.

Several participants were excluded from statistical analyses due to data quality issues. For the N-back task, several participants were excluded due to missing data (*N* = 12), low accuracy (exceeding three standard derivations; *N* = 14), poor brain coverage during MRI scanning (*N* = 21), and excessive motion (*N* = 14). The final sample size for the N-back task was 170 (57.6% female, mean age = 17.02, *SD* = 1.45). For the number-line estimation (NLE) task, data exclusions were due to missing data (*N* = 5), loss of behavioural data (*N* = 4), poor brain coverage (*N* = 21), and excessive motion (*N* = 21). The final sample size for the NLE task was 180 (59.4% female, mean age = 16.90, *SD* = 1.47). For the reading task, data exclusions were also due to missing data (*N* = 10), low accuracy (*N* = 42, whose hit rate <75% and false alarm rate >5%), poor brain coverage (*N* = 23), and excessive motion (*N* = 17). The final sample size for the reading task was 139 (54.7% female, mean age = 17.07, *SD* = 1.49). After quality control, there were 107 participants who had valid data across all three tasks. Exclusion criteria were established prior to lower-level analysis.

Lower-level GLM analyses. For the N-back task, one explanatory variable (EV) modelled the 2-back trials to which a participant correctly responded, with a duration of 2.2 s (equalling the TR). A second EV modelled text cues at the beginning of each block. A third EV modelled 2-back trials with incorrect responses with a duration of 2.2 s. A fourth EV modelled the control (0-back) trials with wrong responses with a duration of 2.2 s. A fifth EV modelled the first two trials in every 2-back block, since the task is meaningful only from the 3rd trial on. A sixth EV modelled two empty TRs after the second trial of the first block and three empty TRs at the end of the run. Therefore, the 0-back control condition was left as the baseline, and the contrast between the 2-back condition and baseline was of primary interest [1 0 0 0 0]. This contrast was designed to isolate higher-level EPs whilst excluding activation engaged by lower-level domain-general processes (e.g., visual item recognition) as the same type of visual displays and stimuli was used in the 2-back and 0-back conditions.

For the NLE task, one EV modelled the task period of a NLE trial which began at the probe display, 1,000 ms before the hatch mark display, and ended at the time of response. Because participants sometimes answered without waiting to process the location of the hatch mark, we considered trials as valid response trials only when participants answered after at least 350 ms. A second EV modelled the same trial time period with a parametric regressor that represented the demeaned base 10 logarithm of the numerical value of the presented stimulus. A third EV modelled the residual time on the trial that corresponds to the time between the participant response and the end of the 2,000 ms period with the hatch mark present on the screen. A fourth EV modelled trials with invalid responses (less than 350 ms or no response recorded). These trials were coded with a duration of 3,000 ms. (corresponding to the 1,000 ms probe period and the 2,000 ms hatch mark period). Similarly, four corresponding EVs modelled the luminance trials. Consequently, what was left in the baseline was the fixation period containing the cue before the probe stimulus, when the line and probe box were present. In the current study, the EV of primary interest was the parametric regressor of the NE trials versus the cue/fixation period.

For the Reading task, one EV modelled the passage block and another EV modelled the word block. A third EV was included to model the two picture-judgement trials. The symbol block was left as the baseline. The EV for the passage against the baseline (symbol reading; 1 0 0) was the primary contrast of interest for this task.

Higher-level GLM analyses. For each lower level contrast, one-sample t-tests, that is a single group average EV in the GLM, were performed at the higher level to obtain the group average of individual activations. Permutation tests are more reliable than parametric tests for neuroimaging data, as they provide better control of false positives and require fewer assumptions (Winkler, Ridgway, Webster, Smith, & Nichols, 2014). The exchangeability assumption of permutation tests requires that the joint distribution of data remains unaltered after shuffling, which means observations need to be independent. However, this independence cannot be guaranteed when familial relationship exists in the sample (e.g., data from twin pairs may be related). Such structured non-independence can be addressed through the usage of exchangeability blocks (EBs), which contain blocks of exchangeable units. Units of data within an EB can be shuffled freely, and EBs can be shuffled as a whole. Hence, we adopted a multi-level block permutation method (Winkler, Webster, Vidaurre, Nichols, & Smith, 2015) implemented in FSL's PALM function (Winkler *et al.*, 2014). In order to account for non-independence due to the family structure of this sample, nested exchangeability blocks were defined that restricted permutations to data with the same family structure (e.g., permutation was allowed between families consisting of two monozygotic twins as a whole, but was not allowed between one family with two monozygotic twins and another family with two dizygotic twins). In this manner, family structure was accounted for without directly modelling these complicated repeated-measures factors. We performed 5,000 permutations for each analysis. We applied a voxel-wise false discovery rate (FDR) correction for multiple comparisons (Genovese, Lazar, & Nichols, 2002). A p value adjusted for FDR across all voxels (size, $2 \times 2 \times 2$ mm) in the brain was produced for each voxel. For main effects of each task, voxels with a 1-p value >.95 (corrected for FDR) were considered significant.

Analysis of site effects. Given our data were acquired in three different fMRI scanners, we examined the possible effects due to site differences. To this end, for each contrast of interest, an additional higher-level GLM was performed. This model contained three explanatory variables (EVs), each corresponding to the lower level input from one scanning site. A site effect was defined as a significant F effect (after FDR correction). To control for potential site differences in statistical power due to differences in sample size across sites, we kept the sample size from each site comparable (N-back task, 69, 52, 49 for three sites or equal [NLE task, 44; reading task, 21]).

Meta-analytic investigation

For the meta-analysis, maps were derived in Neurosynth (http://neurosynth.org), a meta-analytic tool for investigating commonalities in patterns of brain activation across studies (Yarkoni *et al.*, 2011). We utilized the 200-word topics set extracted with a standard topic modelling approach from the abstracts of all articles in the Neurosynth database as of July 2015 (11,406 articles). More specifically, we utilized topic 196 for reading (N = 314 studies; terms are as follows: reading, phonological, word, readers, letter, letters, visual, dyslexia, orthographic, words, processing, dyslexic, children, spelling, rhyming, form, vot, language, strings, developmental, suggest, phonology, consistency, orthography, revealed, dyslexics, va, normal, representations, skill, suggesting, impaired, single, sound, skilled, print, skills, written, decoding, and phoneme), topic 116 for numerical processing (N = 206 studies; terms are as follows: number, numerical, arithmetic, numbers, processing, magnitude, calculation, digit, symbolic, mathematical, distance, math, mental, activation, multiplication, addition, counting, representation, tasks, numerosity, subtraction, retrieval, digits, operations, comparison, quantifiers, quantity, verbal, adults, non-symbolic, arabic, competence, operation, symbols, exact, estimation, single, arrays, quantities, and hips), and topic 022 for EP in WM (N = 485 studies; terms are as follows: memory, working, verbal, term, load, maintenance, performance, information, cognitive, tasks, executive, storage, spatial, capacity, performed, updating, vstm, span, rehearsal, manipulation, function, functions, phonological, cognition, visuospatial, swm, articulatory, ef, sternberg, demands, material, pre, digit, store, networks, delayed, correlates, maintaining, efficient, and increase). Using topics rather than a single word (e.g., 'reading') allow us to capture brain activation associated with a richer set of related concepts (i.e., reading-related processes).

To identify all brain regions associated with each of the reading and math topics, we employed forward inference maps. Such maps identify all regions from visual input to higher-order association areas that are significantly activated during such tasks. For EP, however, we wished to identify those regions *specifically* associated with EP above and beyond other mental processes. To do so, we used a reverse inference map that delineates those brain regions for which there is a significant probability that when activated, a region is associated with the topic of interest above and beyond the average of other topics. However, such maps do not necessarily identify brain regions *uniquely* associated with a given topic, but rather identify those regions that have an increased probability of being activated by tasks that assess the concepts contained within a given topic (e.g., executive) as compared to other terms (e.g., reading).

Conjunction analyses

As there is a difference of opinion on how best to pursue conjunction analyses, we took a middle of the road approach. On the one hand, Nichols, Brett, Andersson, Wager, and Poline (2005) argue that all the comparisons in the conjunction should be individually significant, so that a proper null hypothesis of logical 'AND' is tested. However, Friston, Penny, and Glaser (2005) suggest that such an approach results in a very high threshold that is 'a very conservative procedure, particularly in the context of multiple comparisons' and is 'generally unnecessary'. We made a compromise of between the two viewpoints by using a square root of the significant alpha value (.05) with voxel-wise FDR correction for both three-way and two-way conjunction analyses.

For the empirical analysis, after using FSL's PALM (Winkler, Ridgway, Webster, Smith & Nichols, 2014) to take into account family structure, statistical maps from each task were thresholded at α level of .2236 (square root of .05; FDR corrected). These maps were then binarized. The Fslmaths tool was used to create a map of the set of voxels that activated across both the number-line estimation and reading comprehension tasks at $p < .05$ (.2236 * .2236 = .049). We then set out to determine which of the voxels so identified in this two-way conjunction also showed activation during the performance of the EP (i.e., N-back) task ($p < .011$; i.e., .2236*.2236*.2236), indicative of a three-way conjunction. To determine regions activated for reading comprehension and number-line estimation that did not involve EP, this three-way map was subtracted from the two-way conjunction map of reading comprehension and number-line estimation ($p < .039$; i.e., .2236 * .2236*[1 – .2236] (refer back to Figure 1A).

One might argue that such an approach, whilst setting a threshold that is equivalent across all three tasks, provides a relatively liberal threshold for activation for the EP task (i.e., $p < .2236$) vis a vis the threshold for the conjunction of reading comprehension and number-line estimation ($p < .05$) leading to an overestimation of the contribution of executive regions. Hence, we also computed maps for which the threshold for the EP task ($p < .01$) was equivalent to that for the conjunction of reading comprehension and number-line estimation (.1 for reading and .1 for math yielding $p < .01$ for their conjunction). This conjunction map yielded results of the same pattern as the approach we discuss here (Figure S1).

To identify significant clusters in these conjunction maps, we computed the average Z-value across the relevant maps (e.g., two-way conjunction: reading comprehension and number-line estimation; three-way conjunction: reading comprehension, number-line estimation, and EP) and then masked them with the appropriate binarized map. FSL's cluster command was then used to identify clusters for which the Z value exceeds $p < .05$ after permutation testing and FDR correction.

For the meta-analysis, Neurosynth only provides statistical maps for each topic at a voxel-wise threshold of $p < .01$ with FDR correction. These statistical maps were downloaded and subjected to the same conjunction and cluster identification procedures as the empirical data. This procedure yielded a voxel-wise significance level of $p < .000001$ for the three-way map of reading, math, and EP, and a level of $p < .000099$ for the map of co-activation of reading and math excluding regions activated for EP. The use of a more stringent statistical threshold for the Neurosynth maps as compared to our empirical maps is reasonable given the greater power provided by the multitude of studies in the Neurosynth database.

Results

Empirical data

Behavioural results

All tasks yielded acceptable levels of performance. In the N-back task, participants responded more quickly, $t(169) = -8.94$, $p < .001$, and more accurately, $t(169) = 7.07$, $p < .001$, in the 0-back condition (mean RT = 565 ms, mean accuracy = 94.1%) than in the 2-back condition (mean RT = 628 ms, mean accuracy = 88.9%).

Signal detection theory (d') was used to characterize behavioural performance in the number-line estimation (NLE) task. Here, d' was the difference between the Z scores of hit rate, the ratio of correctly responded trials in the linear trial condition, and the Z scores of false alarm rate, the ratio of trials where the participant indicated the marker was correctly placed in the log or log/linear conditions. The d' value of the number-line estimation condition ($d' = 2.26$) was significantly different from zero, $t(179) = 43.39$, $p < .001$, suggesting that participants were able to linearly map the number to the number axis.

For the reading task, the average hit rate for repeated stimuli was 94% ($SD = 5.8\%$), and average false alarm rate was 0.5% ($SD = 0.7\%$). The average accuracy for picture judging trials after each passage was 88.2% ($SD = 2.3\%$).

Overall group imaging results

Patterns of activation for each of the three tasks were consistent with expectations. For the reading task, the contrast of passages versus symbols yielded large left-sided activation in temporal regions as well as frontal activation (See Figure 3 and Table S1). For the number line task, there was extensive activation across a large portion of occipital cortex as well as the intraparietal sulcus implicated in numerical estimation (see Figure 3 and Table S2). In the N-back task, there was robust activation across portions of the fronto-parietal network and the cingulo-opercular network for the contrast of 2-back versus the – back control condition (see Figure 3 and Table S3). No significant site-related effects, as assessed by an F test, were observed.

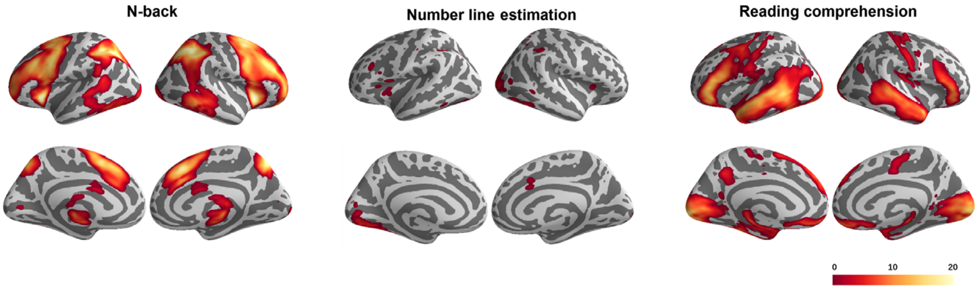


Figure 3. Main effects of the three tasks. The threshold was $p < .05$ (permutation testing, FDR corrected).

Common activation across reading comprehension, number-line estimation, and executive processes in working memory

The three-way conjunction analysis yielded activations in regions of the left middle and inferior frontal gyrus (see Figure 4A, areas in red and Table 1). In addition, regions of the occipitotemporal cortex of the left hemisphere, posterior middle temporal gyrus of the right hemisphere, and cerebellum were also activated by all three tasks. There are 2,014 voxels in this three-way conjunction map, accounting for 3.1% of activation in the N-back results map (when applying a common threshold of $p < .2236$), 3.5% of activation in the reading task (threshold of $p < .2236$), and 17.4% of activation in the NLE task ($p < .2236$).

Common activation across reading comprehension and number-line estimation, excluding regions involved in executive processes in working memory (N-back)

As expected, there was significant overlap in activation between the reading and number-line estimation task that did not involve EP. Most of these were observed in posterior regions of the brain, more specifically occipitotemporal regions (see Figure 4A, areas in blue and Table 2).

Neurosynth data

Common activation across reading, math, and executive processes in working memory

Activations across all three topics were observed in left frontal cortex extending from left precentral gyrus to dorsolateral prefrontal cortex to inferior frontal gyrus. Medial

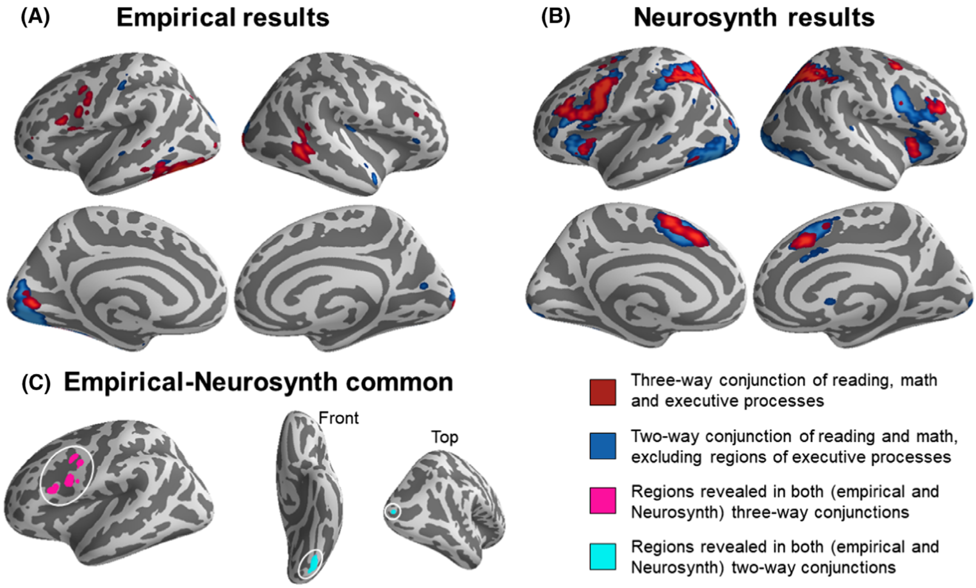


Figure 4. Results of the conjunction analyses across approaches. Both (A) Empirical and (B) Neurosynth results are shown. Regions supporting domain-general executive processes (three-way conjunction) are shown in red. Regions supporting common lower-level functions for math and reading (excluding executive processes) are shown in blue. (C) Regions that yield activation in the both the empirical and Neurosynth analyses. Regions that show activation across all three tasks are shown in magenta, whilst those that activate for both reading and math (but not executive processes) are shown in cyan (D).

Table 1. Clusters that significantly activated empirically across all three tasks: reading comprehension, number-line estimation, and executive processes in working memory (N-back)

Region	Size	BA	*Z*	*X*	*Y*	*Z*
Left Temporal Occipital Fusiform Cortex	1,151	37	6.1	−44	−58	−14
Right Occipital Pole	198	18	6.72	20	−96	0
Right Middle Temporal Gyrus: Temporooccipital part	154	21	5.26	64	−42	−6
Left Inferior Frontal Gyrus: pars opercularis	103	44	7.91	−48	16	26
Left Intracalcarine Cortex	62	17	4.66	−8	−78	8
Right Cerebellum Crus 2	51		6.59	10	−78	−32
Right Middle Temporal Gyrus: Posterior division	50	22	3.46	54	−40	6
Left Middle Frontal Gyrus	47	44	6.46	−50	10	34
Right Middle Frontal Gyrus	40	9	5.65	50	18	30
Left Inferior Frontal Gyrus: pars opercularis	19	44	5.22	−56	14	8
Left Precentral Gyrus	15	44	6.37	−44	4	26
Left Inferior Frontal Gyrus: pars triangularis	12	45	4.05	−44	46	4
Left Occipital Pole	10	18	3.49	−30	−92	2

BA = Brodmann area.

The voxel-wise significance for the three-way conjunction analysis was $p < .011$ (FDR corrected). Clusters with a size of 10 voxels or larger are reported. *Z* values were averaged across the three tasks. The peak with the highest *Z* value within a cluster was reported. *X*, *Y*, *Z*: peak coordinate in MNI space.

Table 2. Clusters that significantly activated empirically for the reading comprehension and number-line estimation tasks, but not executive processes in working memory (N-back)

Region	Size	BA	*Z*	*X*	*Y*	*Z*
Left Occipital Fusiform Gyrus	1,556	17	7.59	−14	−86	−12
Right Occipital Pole	258	18	6.41	22	−96	18
Left Occipital Pole	125	18	4.56	−20	−94	22
Left Postcentral Gyrus	37	4	3.44	−58	−16	48
Right Putamen	21		2.76	−28	−10	0
Right Occipital Fusiform Gyrus	19	19	2.93	32	−72	−16
Right Temporal Pole	19	38	6.51	52	12	−20
Right Intracalcarine Cortex	18	17	4.16	14	−62	10
Right Lateral Occipital Cortex: Superior division	14	37	2.56	50	−66	22
Right Temporal Occipital Fusiform Cortex	13	37	2.51	48	−50	−26
Right Central Opercular Cortex	11	43	3.74	38	−18	18
Right Middle Temporal Gyrus: Temporooccipital part	11	21	3.26	66	−48	2
Right Temporal Pole	11	38	3.93	60	12	−12
Right Hippocampus	11	34	4.35	20	−4	−16

BA = Brodmann area.

The voxel-wise significance threshold was $p < .038$ (FDR corrected). Clusters with a size of 10 voxels or larger are reported. *Z* values were averaged across the reading and number-line estimation tasks. The peak with the highest *Z* value within a cluster was reported. *X*, *Y*, *Z*: peak coordinate in MNI space.

activation in frontal cortex was observed bilaterally in the anterior cingulate cortex and the supplementary motor area (SMA). In addition, activation across all three topics was observed in the superior and inferior parietal lobules bilaterally (see Figure 4B, areas in red, and Table 3). The three-way conjunction map has 2,964 voxels, accounting for 41.5% of activation in the EP topic map, 15.3% of activation in the reading topic map, and 18.6% of activation in the math topic map.

Common Activation across reading and math, excluding regions involved in the executive processes in working memory

Activation was observed across a large number of regions including the left superior parietal lobule, precentral gyrus, bilateral supramarginal gyrus, precentral gyrus, middle frontal gyrus, IFG, paracingulate gyrus, SMA, anterior insula (AI), superior occipital gyrus, inferior temporal gyrus, and lateral occipital cortex (see Figure 4B, areas in blue and Table 4).

Commonalities between the empirical and Neurosynth results

To determine the convergence across our approaches, we identified regions that showed an effect in both the empirical and Neurosynth data (see Figure 4C).

Common activation across reading, math, and executive processes in working memory

This analysis yielded significant activation in left prefrontal regions including the inferior frontal junction (IFJ), premotor regions, and ventral lateral prefrontal cortex (see magenta regions in Figure 4C and Table 5), suggesting these regions may play a critical role in the EP that underlie both reading and math.

Common activation across reading and math, excluding regions involved in the executive processes in working memory

Only one region – the left fusiform gyrus – yielded activation across both approaches (see Figure 4C, areas in cyan, and Table 6). This finding suggests common reliance of reading

Table 3. Clusters that significantly activated in Neurosynth across all three topics: Reading, math, and executive processes

Region	Size	BA	*Z*	*X*	*Y*	*Z*
Left Precentral Gyrus	792	44	19.6	−46	8	28
Left Superior Parietal Lobule	719	7	15.3	−30	−58	44
Left Paracingulate Gyrus	547	6	17.5	−4	10	52
Right Superior Parietal Lobule	534	7	14.9	30	−56	44
Right Middle Frontal Gyrus	95	6	8.57	30	−2	52
Right Insular Cortex	70	47	15.6	34	22	0
Right Middle Frontal Gyrus	60	45	7.14	46	34	24
Left Middle Frontal Gyrus	48	6	7.76	−28	−2	52
Left Insular Cortex	48	6	13.3	−32	22	0
Left Cerebellum Crus1	17		4.62	−40	−62	−28
Right Cerebellum 6	10	37	4.84	28	−60	−28

BA = Brodmann area.

The voxel-wise significance level was $p < .000001$ (FDR corrected). Clusters with a size of 10 voxels or larger are reported. *Z* values were averaged across the three Neurosynth topic maps. The peak with the highest *Z* value within a cluster was reported. *X*, *Y*, *Z*: peak coordinate in MNI space.

Table 4. Clusters that significantly activated in Neurosynth for the topics of reading and math, but not executive processes in working memory

Label	Size	BA	Z	X	Y	Z
Left Superior Parietal Lobule	1,043	7	18.8	−26	−58	46
Left Temporal Occipital Fusiform Cortex	802	37	17.8	−44	−58	−16
Left Superior Frontal Gyrus	716	6	18.8	−4	12	54
Left Middle Frontal Gyrus	675	44	21.9	−48	10	32
Right Lateral Occipital Cortex: superior division	629	7	15.5	30	−60	50
Right Precentral Gyrus	614	44	17.6	46	8	28
Right Inferior Temporal Gyrus: Temporooccipital part	486	37	9.89	46	−60	−8
Left Insular Cortex	409	47	17.9	−34	24	0
Right Insular Cortex	336	47	19.1	34	24	−4
Left Pallidum	54		6.3	−22	8	4
Left Lateral Occipital Cortex: Superior division	37	39	6.04	−50	−62	24
Left Inferior Frontal Gyrus: Pars triangularis	28	45	7.13	−44	38	4
Left Precentral Gyrus	21	6	8.03	−26	−6	52
Left Lateral Occipital Cortex: superior division	19	18	4.74	−30	−86	10
Left Occipital Pole	14	18	4.42	−24	−94	6

BA = Brodmann area.
The voxel-wise significance level for this conjunction analysis was $p < .000099$ (FDR corrected). Clusters with a size larger than 10 were reported. *Z* values were averaged across the reading and math topic maps. The peak with the highest *Z* value within a cluster was reported. *X*, *Y*, *Z*: peak coordinate in MNI space.

and math on a common lower-level neural mechanism for the discrimination of visual form.

Comparison of overlapping activation between reading and math including vs. excluding executive regions

A major goal of our study was to determine the degree to which the neural mechanisms underlying reading and math rely on common executive processes as compared to other domain-general processes (e.g., visual form analysis). As another perspective on this question, we determined the proportion of voxels that co-activated for reading and math that also yielded significant activation for executive processes. As can be seen in Table 7,

Table 5. Common activation across the empirical and Neurosynth results for the conjunction of reading, math, and executive processes in working memory

Label	Size	BA	Z	X	Y	Z
Left Inferior frontal junction	38	44	11.9	−46	12	28
Left Inferior Frontal Gyrus: pars opercularis	15	44	10.1	−46	4	26
Left precentral gyrus	7	6	7.54	−42	2	36
Left precentral gyrus	5	6	7.61	−50	4	36

BA = Brodmann area.
Z values were averaged across all relevant Neurosynth topic maps and empirical maps. The peak with the highest *Z* value within a cluster was reported. *X*, *Y*, *Z*: peak coordinate in MNI space.

Table 6. Common activation across the empirical and Neurosynth results for the 2-way conjunction of reading and math excluding executive regions

Label	Size	BA	*Z*	*X*	*Y*	*Z*
Left fusiform gyrus	28	19	6.89	−36	−70	−14

BA = Brodmann area.
Z values were averaged across all relevant Neurosynth topic maps and empirical maps. The peak with the highest *Z* value within a cluster was reported. *X*, *Y*, *Z*: peak coordinate in MNI space.

of all the voxels that are jointly activated by reading and math (Total), 47.5% in the empirical maps and 33.2% in the Neurosynth maps are accounted for by brain regions that are also activated by executive processes, suggesting that of all domain-general processes that are activated across reading and math, executive processes account for a substantial proportion.

Discussion

Empirical neuroimaging data across multiple tasks measured in the same individuals as well as a meta-analytic approach using the Neurosynth database suggests that executive processes supported by inferior and posterior parts of left lateral prefrontal cortex are important for both reading and math performance. This finding is consistent with the hypothesis that executive operations, which act on information in WM, constitute some of the commonality in neural processing between reading and math. Whilst regions involved in EP are not the only brain regions in which overlapping activation is observed between reading and math tasks, they do account for a substantial proportion of the observed overlap in activation (Table 7). To our knowledge, this is the first study to examine the neural correlates of the overlap between these different dimensions of academic functioning (i.e., reading and math) in the same study.

It is worth commenting on all the potential functions of the lateral prefrontal regions that were revealed across both of our methods. Whilst one must be cautious in inferring psychological processes from patterns of brain activation (Poldrack, 2006), the location of activation may nonetheless provide some insights into the processes involved. One region, premotor cortex, has been linked to articulatory and motoric aspects of WM (Price, 2012). Another – the inferior frontal junction (IFJ) – has been proposed to be an important region for EP (Derrfuss, Brass, Neumann, & von Cramon, 2005), including selecting task-relevant information, and switching between tasks or representational sets (Derrfuss *et al.*, 2005). The final region – ventrolateral prefrontal cortex (VLPFC) including Broca's area (BA44) – has traditionally been viewed as a region involved in language production (Hagoort, 2014), comprehension (e.g., Rodd, Davis, & Johnsrude,

Table 7. Number of voxels contained in overlap maps for reading and math

	Activation across reading, math and executive processes	Activation across reading and math but not executive processes	Total
Empirical	2,014 (47.5%)	2,227 (52.5%)	4,241
Neurosynth	2,964 (33.2%)	5,964 (66.8%)	8,928

2005), and syntax (e.g., Vigneau *et al.*, 2006). However, this region also activates during math tasks such as exact arithmetic (Arsalidou & Taylor, 2011) and mathematical logic (Houde & Tzourio-Mazoyer, 2003). Our findings are consistent with other work suggesting that this region may have a more domain-general executive role in resolving interference amongst competing items in WM (Nelson, Reuter-Lorenz, Persson, Sylvester, & Jonides, 2009), helping to select and retrieve goal-relevant information from long-term memory (Hirshorn & Thompson-Schill, 2006) or organizing information in a hierarchical manner (Bahlmann, Schubotz, Mueller, Koester, & Friederici, 2009).

Considering the various functions of these three regions, it seems that multiple types, rather than a single type, of EPs are involved in reading and math. This observation is consistent with a recent meta-analysis on behavioural studies, which suggests that different subtypes of EP almost indistinguishably contribute to reading and math (Jacob & Parkinson, 2015). However, there may be a hierarchy of these EP functions. At a higher level, IFJ and VLPFC may work together to actively select task-relevant information. Such processes are exerted upon information from different cognitive domains, including language and math. Then, at a lower level, the premotor cortex may be involved in determining or executing the specific actions consistent with those higher-level processes.

It is worth noting that the main lateral prefrontal region for which we did not observe a three-way conjunction of activation across both the empirical and meta-analytic approaches was mid-DLPFC. In many models, this region serves to bias information processing towards goal-relevant information (Banich, 2009; Miller & Cohen, 2001). This region has also been implicated in successful maintenance of relevant information in the presence of distraction (Feredoes, Heinen, Weiskopf, Ruff, & Driver, 2011) and/or in the face of pre-potent but task-irrelevant information (Snyder, Banich, & Munakata, 2014). The reasons why this region did not show activation in our study are not clear, but there are a number of possibilities. First, the lack of activation in these regions may have occurred because these regions are still developing with regard to executive processes (e.g., Andrews-Hanna *et al.*, 2011) in our adolescent sample. Another possibility is that portions of mid-DLPFC function are activated in a more highly domain-specific way at this age, as different portions of DLPFC were activatied in the main effect of the three tasks (Figure 3). Finally, the specific reading comprehension and numerical estimation paradigms we employed in our empirical investigation may not tap this type of control process, although it may in fact be engaged across more demanding reading and math tasks. One way to more directly address this issue in future studies might be to design a study that additively manipulates reading, math, and executive operations within the same paradigm.

What then is to be made of the finding that most regions of lateral prefrontal cortex show overlap in activation across all three domains: reading, math & EP in WM? We propose that whilst certain regions of the left hemisphere, such as Broca's area, have been conceptualized as 'language related', they may potentially be better conceptualized as supporting domain-general EP that are utilized during both language processing and other higher-order cognitive operations that involve systematic relationships amongst items. For example, sentence comprehension requires an understanding of the relationship between words within a syntactic frame, and mathematical abilities involving magnitude estimation require mapping of values onto a reference frame of the number line.

One difference between results of empirical and Neurosynth analyses is that left inferior occipitotemporal regions were activated in all three empirical tasks but were activated in only two Neurosynth topic (reading and math) maps (Figure 4A, B). This area is located in the ventral stream of visual processing, which is related to object recognition and formation of representations (Goodale & Milner, 1992; Mishkin, Ungerleider, & Macko, 1983). Since all empirical tasks are visual tasks, activation in this area is expected. On the other hand, as for the Neurosynth analysis, we chose the reverse inference map for the EP topic, which by definition controls for the modalities of the stimuli on which EP acts, which likely accounts for the absence of occipitotemporal activation in this map. We suggest that the regions highlighted in both empirical and Neurosynth analyses may best represent neural substrates of domain-general EP (Figure 4C), as they took advantage of these two complementary methods.

Clinical implications

In addition to the important implications for cognitive models of executive processes as they influence academic achievement, the current results suggest that future studies should test whether a shared weakness in EP mediated by left lateral prefrontal cortex may help to explain the frequent co-occurrence between clinical diagnoses of reading disability and math disability. Understanding the causes of the ubiquitous comorbidity amongst developmental disorders such as learning disabilities, attention-deficit/hyperactivity disorder, and autism spectrum disorders is a key direction for future research, and the current results illustrate the potential utility of the design described in this paper to be used as a tool to address this question. Such design should be used in combination with emerging techniques like machine learning and multimodality approaches (Astle, Bathelt, Team, & Holmes, 2019; Lerman-Sinkoff *et al.*, 2017).

Further, whilst EP account for a significant proportion of the covariance between reading and math, additional covariance remains to be explained. The pattern we observed is consistent with multiple deficit models of reading and math difficulties and suggests that weaknesses in EP are an important domain-general factor in multifactorial neurocognitive models in which weaknesses in several cognitive domains act in combination to contribute to learning disabilities (Pennington, 2006; Willcutt *et al.*, 2013). The major brain region not involved in EP in WM, but which nonetheless that showed consistent overlap in activation between reading and math was the ventral visual processing stream. This overlap may simply reflect the fact that both math and reading rely upon visual input and symbols that are subsequently translated to more abstract representations with amodal meaning. Exactly how executive processes might influence the learning of reading and math during this adolescent developmental stage of life is not clear, although this issue has been considered with regard to pre-school and grade-school children (Church *et al.*, 2019; Cragg *et al.*, 2017).

Limitations

In interpreting our results, it is important to note that the NLE task yielded smaller regions of activation than those of the N-back and reading tasks (Figure 4). We hold that this may be due to a difference in task design: the NLE task utilized an event-related design whilst the other two tasks used a block design. Therefore, although the task itself was able to adequately assess activation with numerical estimation, this difference in the design across tasks may have caused the NLE task to be relatively under-powered compared to the other two tasks. As such, it might be that if a block design was employed to assess numerical estimation, one might find additional EP regions beyond those we observed that would show overlap in activation across all three tasks.

One should consider the limitations imposed by the scope of empirical tasks we used. The number line task taps a rather basic aspect of math ability, that is number sense and magnitude processing. Nonetheless, it should be noted that performance on magnitude processing is highly predictive of mathematical competence and that this relationship does not reduce substantially with age (Schneider *et al.*, 2017). In addition, the reading task was also not especially challenging as the sentences, although topical across frames, did not involve higher-order integration across paragraphs. From this perspective, our results most likely downplay the role at EP may play in more advanced reading and mathematical skills. The results from the meta-analytic platform, Neurosynth, are based on text analysis of published studies and as such do not provide information on the types of tasks nor the difficulty of the studies contributing to its results. Nonetheless, the topics do cover a broad range of processes underlying each of reading and of math. It should be noted that the limitation of each of these approaches is in some sense the opposite. The empirical approach was limited to a restricted set of tasks tapping a narrow range of processes involved in math, reading, and executive function. In contrast, the Neurosnyth approach was much broader encompassing a wide variety of tasks that are involved in such processes. As such, the convergence of findings across both approaches is notable.

A final caveat should be noted. The conclusion of both analyses in this study is based on averaged result of a large number of participants and do not take individual differences into account. However, the neural substrates of domain-general EP may vary for different sub-groups of adolescents, such as those of differing or levels of ability. Since this study was not designed to examine individual differences in EP, this would be an interesting topic for a future research.

Conclusion and implications

Interpretations of our results must be put in the context of the developmental status of our participants, whose average age was approximately seventeen. At this point in their lives, our participants would likely be relatively proficient in their comprehension of the simple paragraphs utilized in our study, as well as the number sense tapped by the number line task (Chall, 1983a; Siegler & Opfer, 2003). As has been suggested by others, it is critical to focus on domain-specific skills during grade school, such as phonological processing in reading (Wagner & Torgesen, 1987) and number sense in math (Namkung & Fuchs, 2016) that serve as a foundation upon which more complex skills can develop. Once mastery over such basic underlying processes (e.g., number sense, phonology, and vocabulary) has been established, more sophisticated use of language and mathematical reasoning may be enabled by frontal lobe-supported domain-general EP in WM that allow such information to be accessed and deployed effectively. Taken together, the results suggest that the associations observed between reading and math may result from their mutual reliance on EP in WM, which are supported by the posterior and lateral parts of left PFC that act on and enable the effective use of information in WM.

Acknowledgements

Research reported in this paper was supported by the Eunice Kennedy Shriver National Institute of Child Health and Human Development (NICHD) of the National Institutes of Health under award number HD075460, and the authors were also supported by NICHD Learning Disabilities Research Center grant (P50 HD 27802). Data entry via the REDCap system was supposed by grant UL1TR001070 from the National Center For Advancing Translational Sciences. DRL is also supported by NICHD F31HD091967. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. We thank Kathy Pearson for her computational help with data archiving, quality assurance, and pre-processing.

Conflicts of interest

All authors declare no conflict of interest.

Author contributions

Kai Wang, Ph.D. (Formal analysis; Investigation; Methodology; Software; Visualization; Writing – original draft). Marie T. Banich (Conceptualization; Funding acquisition; Methodology; Project administration; Supervision; Writing – original draft; Writing – review & editing). Andrew E. Reineberg (Data curation; Formal analysis; Investigation; Writing – review & editing). Daniel R. Leopold (Data curation; Writing – review & editing). Erik G. Willcutt (Conceptualization; Funding acquisition; Writing – review & editing). Laurie E. Cutting (Conceptualization; Formal analysis; Funding acquisition; Writing – review & editing). Stephanie N. Del Tufo (Writing – review & editing). Lee A. Thompson (Conceptualization; Funding acquisition; Writing – review & editing). John Opfer (Data curation; Formal analysis). Frank J. Kanayet (Formal analysis). Zhong-Lin Lu (Conceptualization; Funding acquisition; Writing – review & editing). Stephen A. Petrill (Conceptualization; Funding acquisition; Project administration; Writing – review & editing).

Data availability statement

The data that support the findings of this study are available on request from the corresponding author, (K. W). The data are not publicly available as they contain information that may compromise confidentiality and informed consent of participants.

References

Aboud, K. S., Bailey, S. K., Petrill, S. A., & Cutting, L. E. (2016). Comprehending text versus reading words in young readers with varying reading ability: Distinct patterns of functional connectivity from common processing hubs. *Developmental Science*, *19*, 632–656. https://doi.org/10.1111/desc.12422

Andrews-Hanna, J. R., Seghete, K. L. M., Claus, E. D., Burgess, G. C., Ruzic, L., & Banich, M. T. (2011). Cognitive control in adolescence: Neural underpinnings and relation to self-report behaviors. *PLoS ONE*, *6*(6), e21598. https://doi.org/10.1371/journal.pone.0021598

Ansari, D., Fugelsang, J. A., Dhital, B., & Venkatraman, V. (2006). Dissociating response conflict from numerical magnitude processing in the brain: An event-related fMRI study. *NeuroImage*, *32*, 799–805. https://doi.org/10.1016/j.neuroimage.2006.04.184

Aron, A. R., Robbins, T. W., & Poldrack, R. A. (2004). Inhibition and the right inferior frontal cortex. *Trends in Cognitive Sciences*, *8*, 170–177. https://doi.org/10.1016/j.tics.2004.02.010

Arsalidou, M., & Taylor, M. J. (2011). Is 2+2=4? Meta-analyses of brain areas needed for numbers and calculations. *NeuroImage*, *54*, 2382–2393. https://doi.org/10.1016/j.neuroimage.2010.10.009

Astle, D. E., Bathelt, J., Team, T. C., & Holmes, J. (2019). Remapping the cognitive and neural profiles of children who struggle at school. *Developmental Science*, *22*(1), e12747. https://doi.org/10.1111/desc.12747

Badre, D., & Wagner, A. D. (2007). Left ventrolateral prefrontal cortex and the cognitive control of memory. *Neuropsychologia*, *45*, 2883–2901. https://doi.org/10.1016/j.neuropsychologia.2007.06.015

Bahlmann, J., Schubotz, R. I., Mueller, J. L., Koester, D., & Friederici, A. D. (2009). Neural circuits of hierarchical visuo-spatial sequence processing. *Brain Research*, *1298*, 161–170. https://doi.org/10.1016/j.brainres.2009.08.017

Banich, M. T. (2009). Executive function: The search for an integrated account. *Current Directions in Psychological Science*, *18*, 89–94. https://doi.org/10.1111/j.1467-8721.2009.01615.x

Barbey, A. K., Koenigs, M., & Grafman, J. (2013). Dorsolateral prefrontal contributions to human working memory. *Cortex*, *49*, 1195–1205. https://doi.org/10.1016/j.cortex.2012.05.022

Best, J. R., & Miller, P. H. (2010). A developmental perspective on executive function. *Child Development*, *81*, 1641–1660. https://doi.org/10.1111/j.1467-8624.2010.01499.x

Best, J. R., Miller, P. H., & Naglieri, J. A. (2011). Relations between executive function and academic achievement from ages 5 to 17 in a large, representative national sample. *Learning and Individual Differences*, *21*(4), 327–336. https://doi.org/10.1016/j.lindif.2011.01.007

Biederman, J., Monuteaux, M. C., Doyle, A. E., Seidman, L. J., Wilens, T. E., Ferrero, F., . . . Faraone, S. V. (2004). Impact of executive function deficits and attention-deficit/hyperactivity disorder (ADHD) on academic outcomes in children. *Journal of Consulting and Clinical Psychology*, *72*, 757–766. https://doi.org/10.1037/0022-006X.72.5.757

Bull, R., Espy, K. A., & Wiebe, S. A. (2008). Short-term memory, working memory, and executive functioning in preschoolers: Longitudinal predictors of mathematical achievement at age 7 years. *Developmental Neuropsychology*, *33*(3), 205–228. https://doi.org/10.1080/87565640801982312

Bull, R., & Lee, K. (2014). Executive functioning and mathematics achievement. *Child Development Perspectives*, *8*(1), 36–41. https://doi.org/10.1111/cdep.12059

Casey, B. J., Jones, R. M., & Hare, T. A. (2008). The Adolescent Brain. *Annals of the New York Academy of Sciences*, *1124*(1), 111–126. https://doi.org/10.1196/annals.1440.010

Chall, J. S. (1983a). *Learning to read: The great debate*. McGraw-Hill.

Chall, J. S. (1983b). *Stages of reading development*. New York, NY: McGraw-Hill.

Christopher, M. E., Hulslander, J., Byrne, B., Samuelsson, S., Keenan, J. M., Pennington, B., . . . Olson, R. K. (2015). Genetic and environmental etiologies of the longitudinal relations between prereading skills and reading. *Child Development*, *86*(2), 342–361. https://doi.org/10.1111/cdev.12295

Church, J. A., Cirino, P. T., Miciak, J., Juranek, J., Vaughn, S., & Fletcher, J. M. (2019). Cognitive, intervention, and neuroimaging perspectives on executive function in children with reading disabilities. *New Directions for Child and Adolescent Development*, *2019*, 25–54. https://doi.org/10.1002/cad.20292

Cirino, P. T., Miciak, J., Ahmed, Y., Barnes, M. A., Taylor, W. P., & Gerst, E. H. (2019). Executive function: Association with multiple reading skills. *Reading and Writing*, *32*, 1819–1846. https://doi.org/10.1007/s11145-018-9923-9

Cole, M. W., Repovš, G., & Anticevic, A. (2014). The Frontoparietal control system: A central role in mental health. *The Neuroscientist*, *20*(6), 652–664. https://doi.org/10.1177/1073858414525995

Cragg, L., Keeble, S., Richardson, S., Roome, H. E., & Gilmore, C. (2017). Direct and indirect influences of executive functions on mathematics achievement. *Cognition*, *162*, 12–26. https://doi.org/10.1016/j.cognition.2017.01.014

Depue, B. E., Orr, J., Smolker, H., Naaz, F., & Banich, M. (2015). The organization of right prefrontal networks reveals common mechanisms of inhibitory regulation across cognitive, emotional, and motor processes. *Cerebral Cortex*, *26*, 1634–1646. https://doi.org/10.1093/cercor/bhu324

Derrfuss, J., Brass, M., Neumann, J., & von Cramon, D. Y. (2005). Involvement of the inferior frontal junction in cognitive control: Meta-analyses of switching and Stroop studies. *Human Brain Mapping*, *25*(1), 22–34. https://doi.org/10.1002/hbm.20127

Diamond, A. (2013). Executive functions. *Annual Review of Psychology*, *64*, 135–168. https://doi.org/10.1146/annurev-psych-113011-143750

Duncan, J. (2010). The multiple-demand (MD) system of the primate brain: Mental programs for intelligent behaviour. *Trends in Cognitive Sciences*, *14*, 172–179. https://doi.org/10.1016/j.tics.2010.01.004

Fedorenko, E., & Thompson-Schill, S. L. (2014). Reworking the language network. *Trends in Cognitive Sciences*, *18*, 120–126. https://doi.org/10.1016/j.tics.2013.12.006

Feredoes, E., Heinen, K., Weiskopf, N., Ruff, C., & Driver, J. (2011). Causal evidence for frontal involvement in memory target maintenance by posterior brain areas during distracter interference of visual working memory. *Proceedings of the National Academy of Sciences*, *108*, 17510–17515. https://doi.org/10.1073/pnas.1106439108

Ferstl, E. C., Neumann, J., Bogler, C., & von Cramon, D. Y. (2008). The extended language network: A meta-analysis of neuroimaging studies on text comprehension. *Human Brain Mapping*, *29*(5), 581–593. https://doi.org/10.1002/hbm.20422

Friedman, N. P., & Miyake, A. (2017). Unity and diversity of executive functions: Individual differences as a window on cognitive structure. *Cortex*, *86*, 186–204. https://doi.org/10.1016/j.cortex.2016.04.023

Friston, K. J., Penny, W. D., & Glaser, D. E. (2005). Conjunction revisited. *NeuroImage*, *25*, 661–667. https://doi.org/10.1016/j.neuroimage.2005.01.013

Gajewski, P. D., Hanisch, E., Falkenstein, M., Thönes, S., & Wascher, E. (2018). What does the n-back task measure as we get older? Relations between working-memory measures and other cognitive functions across the lifespan. *Frontiers in Psychology*, *9*, 2208. https://doi.org/10.3389/fpsyg.2018.02208

Genovese, C. R., Lazar, N. A., & Nichols, T. (2002). Thresholding of statistical maps in functional neuroimaging using the false discovery rate. *NeuroImage*, *15*, 870–878. https://doi.org/10.1006/nimg.2001.1037

Gogtay, N., Giedd, J. N., Lusk, L., Hayashi, K. M., Greenstein, D., Vaituzis, A. C., . . . Thompson, P. M. (2004). Dynamic mapping of human cortical development during childhood through early adulthood. *Proceedings of the National Academy of Sciences of the United States of America*, *101*, 8174–8179. https://doi.org/10.1073/pnas.0402680101

Goodale, M. A., & Milner, A. D. (1992). Separate visual pathways for perception and action. *Trends in Neurosciences*, *15*(1), 20–25. https://doi.org/10.1016/0166-2236(92)90344-8

Hagoort, P. (2014). Nodes and networks in the neural architecture for language: Broca's region and beyond. *Current Opinion in Neurobiology*, *28*, 136–141. https://doi.org/10.1016/j.conb.2014.07.013

Halberda, J., Ly, R., Wilmer, J. B., Naiman, D. Q., & Germine, L. (2012). Number sense across the lifespan as revealed by a massive Internet-based sample. *Proceedings of the National Academy of Sciences of the United States of America*, *109*, 11116–11120. https://doi.org/10.1073/pnas.1200196109

Hannagan, T., Amedi, A., Cohen, L., Dehaene-Lambertz, G., & Dehaene, S. (2015). Origins of the specialization for letters and numbers in ventral occipitotemporal cortex. *Trends in Cognitive Sciences*, *19*, 374–382. https://doi.org/10.1016/j.tics.2015.05.006

Hirshorn, E. A., & Thompson-Schill, S. L. (2006). Role of the left inferior frontal gyrus in covert word retrieval: Neural correlates of switching during verbal fluency. *Neuropsychologia*, *44*, 2547–2557. https://doi.org/10.1016/j.neuropsychologia.2006.03.035

Houde, O., & Tzourio-Mazoyer, N. (2003). Neural foundations of logical and mathematical cognition. *Nature Reviews Neuroscience*, *4*(6), 507–514. https://doi.org/10.1038/nrn1117

Jacob, R., & Parkinson, J. (2015). The potential for school-based interventions that target executive function to improve academic achievement: A review. *Review of Educational Research*, *85*(4), 512–552. https://doi.org/10.3102/0034654314561338

Jaeggi, S. M., Buschkuehl, M., Perrig, W. J., & Meier, B. (2010). The concurrent validity of the N-back task as a working memory measure. *Memory*, *18*(4), 394–412. https://doi.org/10.1080/09658211003702171

January, D., Trueswell, J. C., & Thompson-Schill, S. L. (2009). Co-localization of Stroop and syntactic ambiguity resolution in Broca's area: Implications for the neural basis of sentence processing. *Journal of Cognitive Neuroscience*, *21*, 2434–2444. https://doi.org/10.1162/jocn.2008.21179

Jenkinson, M., Bannister, P., Brady, M., & Smith, S. (2002). Improved optimization for the robust and accurate linear registration and motion correction of brain images. *NeuroImage*, *17*(2), 825–841. https://doi.org/10.1006/nimg.2002.1132

Kane, M. J., Conway, A. R., Miura, T. K., & Colflesh, G. J. (2007). Working memory, attention control, and the N-back task: A question of construct validity. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *33*, 615–622. https://doi.org/10.1037/0278-7393.33.3.615

Lan, X., Legare, C. H., Ponitz, C. C., Li, S., & Morrison, F. J. (2011). Investigating the links between the subcomponents of executive function and academic achievement: A cross-cultural analysis of Chinese and American preschoolers. *Journal of Experimental Child Psychology*, *108*, 677–692. https://doi.org/10.1016/j.jecp.2010.11.001

Lerman-Sinkoff, D. B., Sui, J., Rachakonda, S., Kandala, S., Calhoun, V. D., & Barch, D. M. (2017). Multimodal neural correlates of cognitive control in the Human Connectome Project. *NeuroImage*, *163*, 41–54. https://doi.org/10.1016/j.neuroimage.2017.08.081

Mencarelli, L., Neri, F., Momi, D., Menardi, A., Rossi, S., Rossi, A., & Santarnecchi, E. (2019). Stimuli, presentation modality, and load-specific brain activity patterns during n-back task. *Human Brain Mapping*, *40*, 3810–3831.

Miller, E. K., & Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience*, *24*(1), 167–202. https://doi.org/10.1146/annurev.neuro.24.1.167

Mishkin, M., Ungerleider, L. G., & Macko, K. A. (1983). Object vision and spatial vision: Two cortical pathways. *Trends in Neurosciences*, *6*, 414–417. https://doi.org/10.1016/0166-2236(83)90190-X

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex "Frontal Lobe" tasks: A latent variable analysis. *Cognitive Psychology*, *41*(1), 49–100. https://doi.org/10.1006/cogp.1999.0734

Muhle-Karbe, P. S., Derrfuss, J., Lynn, M. T., Neubert, F. X., Fox, P. T., Brass, M., & Eickhoff, S. B. (2016). Co-activation-based parcellation of the lateral prefrontal cortex delineates the inferior frontal junction area. *Cerebral Cortex*, *26*, 2225–2241. https://doi.org/10.1093/cercor/bhv073

Namkung, J. M., & Fuchs, L. S. (2016). Cognitive predictors of calculations and number line estimation with whole numbers and fractions among at-risk students. *Journal of Educational Psychology*, *108*(2), 214–228. https://doi.org/10.1037/edu0000055

Nelson, J. K., Reuter-Lorenz, P. A., Persson, J., Sylvester, C.-Y. C., & Jonides, J. (2009). Mapping interference resolution across task domains: A shared control process in left inferior frontal gyrus. *Brain Research*, *1256*, 92–100. https://doi.org/10.1016/j.brainres.2008.12.001

Nichols, T., Brett, M., Andersson, J., Wager, T., & Poline, J. B. (2005). Valid conjunction inference with the minimum statistic. *NeuroImage*, *25*, 653–660. https://doi.org/10.1016/j.neuroimage.2004.12.005

Nieder, A. (2016). The neuronal code for number. *Nature Reviews Neuroscience*, *17*, 366–382. https://doi.org/10.1038/nrn.2016.40

Pennington, B. F. (2006). From single to multiple deficit models of developmental disorders. *Cognition*, *101*(2), 385–413. https://doi.org/10.1016/j.cognition.2006.04.008

Petrides, M. (2000). The role of the mid-dorsolateral prefrontal cortex in working memory. *Experimental Brain Research*, *133*(1), 44–54. https://doi.org/10.1007/s002210000399

Poldrack, R. A. (2006). Can cognitive processes be inferred from neuroimaging data? *Trends in Cognitive Sciences*, *10*(2), 59–63. https://doi.org/10.1016/j.tics.2005.12.004

Poldrack, R. A., Mumford, J. A., & Nichols, T. E. (2011). *Preprocessing fMRI data. Handbook of functional MRI data analysis* (pp. 34–50). New York, NY: Cambridge University Press.

Price, C. J. (2012). A review and synthesis of the first 20 years of PET and fMRI studies of heard speech, spoken language and reading. *NeuroImage*, *62*, 816–847. https://doi.org/10.1016/j.neuroimage.2012.04.062

Rac-Lubashevsky, R., & Kessler, Y. (2016). Decomposing the n-back task: An individual differences study using the reference-back paradigm. *Neuropsychologia*, *90*, 190–199. https://doi.org/10.1016/j.neuropsychologia.2016.07.013

Rajah, M. N., Ames, B., & D'Esposito, M. (2008). Prefrontal contributions to domain-general executive control processes during temporal context retrieval. *Neuropsychologia*, *46*, 1088–1103. https://doi.org/10.1016/j.neuropsychologia.2007.10.023

Rodd, J. M., Davis, M. H., & Johnsrude, I. S. (2005). The neural mechanisms of speech comprehension: fMRI studies of semantic ambiguity. *Cerebral Cortex*, *15*, 1261–1269. https://doi.org/10.1093/cercor/bhi009

Rose, S. A., Feldman, J. F., & Jankowski, J. J. (2011). Modeling a cascade of effects: The role of speed and executive functioning in preterm/full-term differences in academic achievement. *Developmental Science*, *14*, 1161–1175. https://doi.org/10.1111/j.1467-7687.2011.01068.x

Schneider, M., Beeres, K., Coban, L., Merz, S., Susan Schmidt, S., Stricker, J., & De Smedt, B. (2017). Associations of non-symbolic and symbolic numerical magnitude processing with mathematical competence: A meta-analysis. *Developmental Science*, *20*(3), e12372. https://doi.org/10.1111/desc.12372

Seghier, M. L. (2013). The angular gyrus: Multiple functions and multiple subdivisions. *Neuroscientist*, *19*(1), 43–61. https://doi.org/10.1177/1073858412440596

Siegler, R. S., & Opfer, J. E. (2003). The development of numerical estimation: Evidence for multiple representations of numerical quantity. *Psychological Science*, *14*(3), 237–250. https://doi.org/10.1111/1467-9280.02438

Simmons, W. K., & Martin, A. (2009). The anterior temporal lobes and the functional architecture of semantic memory. *Journal of the International Neuropsychological Society*, *15*, 645–649. https://doi.org/10.1017/S1355617709990348

Snyder, H. R., Banich, M. T., & Munakata, Y. (2014). All Competition is not alike: Neural mechanisms for resolving underdetermined and prepotent competition. *Journal of Cognitive Neuroscience*, *26*, 2608–2623. https://doi.org/10.1162/jocn_a_00652

Snyder, H. R., Hutchison, N., Nyhus, E., Curran, T., Banich, M. T., O'Reilly, R. C., & Munakata, Y. (2010). Neural inhibition enables selection during language processing. *Proceedings of the National Academy of Sciences*, *107*(38), 16483–16488.

Soden, B., Christopher, M. E., Hulslander, J., Olson, R. K., Cutting, L., Keenan, J. M., . . . Petrill, S. A. (2015). Longitudinal stability in reading comprehension is largely heritable from grades 1 to 6. *PLoS ONE*, *10*(1), e0113807. https://doi.org/10.1371/journal.pone.0113807

Vigneau, M., Beaucousin, V., Hervé, P. Y., Duffau, H., Crivello, F., Houdé, O., . . . Tzourio-Mazoyer, N. (2006). Meta-analyzing left hemisphere language areas: Phonology, semantics, and sentence processing. *NeuroImage*, *30*, 1414–1432. https://doi.org/10.1016/j.neuroimage.2005.11.002

Vincent, J. L., Kahn, I., Snyder, A. Z., Raichle, M. E., & Buckner, R. L. (2008). Evidence for a frontoparietal control system revealed by intrinsic functional connectivity. *Journal of Neurophysiology*, *100*, 3328–3342. https://doi.org/10.1152/jn.90355.2008

Wagner, R. K., & Torgesen, J. K. (1987). The nature of phonological processing and its causal role in the acquisition of reading skills. *Psychological Bulletin*, *101*(2), 192. https://doi.org/10.1037/0033-2909.101.2.192

Wang, K., Leopold, D. R., Banich, M. T., Reineberg, A. E., Willcutt, E. G., Cutting, L. E., . . . Petrill, S. A. (2019). Characterizing and decomposing the neural correlates of individual differences in reading ability among adolescents with task-based fMRI. *Developmental Cognitive Neuroscience*, *37*, 100647. https://doi.org/10.1016/j.dcn.2019.100647

Wilkey, E. D., & Price, G. R. (2018). Attention to number: The convergence of numerical magnitude processing, attention, and mathematics in the inferior frontal gyrus. *Human Brain Mapping*, *40*, 928–943. https://doi.org/10.1002/hbm.24422

Willcutt, E. G., Petrill, S. A., Wu, S., Boada, R., DeFries, J. C., Olson, R. K., & Pennington, B. F. (2013). Comorbidity Between reading disability and math disability: Concurrent psychopathology, functional impairment, and neuropsychological functioning. *Journal of learning disabilities*, *46*(6), 500–516. https://doi.org/10.1177/0022219413477476

Winkler, A. M., Ridgway, G. R., Webster, M. A., Smith, S. M., & Nichols, T. E. (2014). Permutation inference for the general linear model. *NeuroImage*, *92*, 381–397. https://doi.org/10.1016/j.neuroimage.2014.01.060

Winkler, A. M., Webster, M. A., Vidaurre, D., Nichols, T. E., & Smith, S. M. (2015). Multi-level block permutation. *NeuroImage*, *123*, 253–268. https://doi.org/10.1016/j.neuroimage.2015.05.092

Yarkoni, T., Poldrack, R. A., Nichols, T. E., Van Essen, D. C., & Wager, T. D. (2011). Large-scale automated synthesis of human functional neuroimaging data. *Nature Methods*, *8*, 665. https://doi.org/10.1038/nmeth.1635

Ye, Z., & Zhou, X. (2009). Conflict control during sentence comprehension: fMRI evidence. *NeuroImage*, *48*(1), 280–290. https://doi.org/10.1016/j.neuroimage.2009.06.032

Yeniad, N., Malda, M., Mesman, J., van Ijzendoorn, M. H., & Pieper, S. (2013). Shifting ability predicts math and reading performance in children: A meta-analytical study. *Learning and Individual Differences*, *23*, 1–9. https://doi.org/10.1016/j.lindif.2012.10.004

Received 1 March 2019; revised version received 16 December 2019

Supporting Information

The following supporting information may be found in the online edition of the article:

Figure S1. Conjunction result using more stringent thresholds.
Table S1. Brain activations for passage reading effect in the reading task.
Table S2. Brain activations for number size parametric effect in the NLE task.
Table S3. Brain activations for 2-back vs. 0-back in the N-back task.