

Shared genetic aetiology between cognitive performance and brain activations in language and math tasks

Authors:

Yann Le Guen^{1}, Marie Amalric², Philippe Pinel², Christophe Pallier², Vincent Frouin^{1*}*

Affiliations:

¹Neurospin, Institut Joliot, CEA, Université Paris-Saclay, Gif-sur-Yvette, France

²Cognitive Neuroimaging Unit, U992, INSERM, Neurospin, Institut Joliot, CEA, Université Paris-Saclay, Gif-sur-Yvette, France

* Corresponding authors' information:

Emails: yann.leguen@cea.fr (YLG), vincent.frouin@cea.fr (VF)

Cognitive performance is highly heritable. However, little is known about common genetic influences on cognitive ability and brain activation when engaged in a cognitive task. The Human Connectome Project (HCP) offers a unique opportunity to study this shared genetic etiology with an extended pedigree of 785 individuals. To investigate this common genetic origin, we took advantage of the HCP dataset, which includes both language and mathematics activation tasks. Using the HCP multimodal parcellation, we identified areals in which inter-individual functional MRI (fMRI) activation variance was significantly explained by genetics. Then, we performed bivariate genetic analyses between the neural activations and behavioral scores, corresponding to the fMRI task accuracies, fluid intelligence, working memory and language performance. We observed that several parts of the language network along the superior temporal sulcus, as well as the angular gyrus belonging to the math processing network, are significantly genetically correlated with these indicators of cognitive performance. This shared genetic etiology provides insights into the brain areas where the human-specific genetic repertoire is expressed. Studying the association of polygenic risk scores, using variants associated with human cognitive ability and brain activation, would provide an opportunity to better understand where these variants are influential.

Introduction

Language and math functions in humans are extensively studied in fundamental neuroscience as distinctive abilities of human lineage. They are frequently assessed through neuroimaging to provide endophenotypes^{1,2}. They are used as a way to classify the broad behavioral symptoms of language impairments into stable phenotypes that in turn are candidates to search for potential associations with either medical treatment responses or genetic profiles^{3,4}. Structural properties observed using magnetic resonance imaging (MRI) or activations observed with functional MRI (fMRI) have been used to produce such endophenotypes⁵. These can reveal differences between control and disease groups in language-specific regions⁶ or distinguish disorder subtypes such as grammatical-SLI (specific language impairment)⁷.

Imaging-genetics resources, such as the Human Connectome Project (HCP) provide an unprecedented opportunity to study the variability of such endophenotypes in control subjects, as well as to determine their potential heritability or association with genetics. Following up on these ideas, we first proposed to study the additive genetic variance involved in fMRI activation differences among typically developed individuals. We used the pedigree data from the HCP language comprehension and verbal math fMRI tasks. These tasks recruit regions directly implicated in brain disorders, such as Broca's area in SLI⁸, the angular gyrus in developmental dyslexia¹ and the intraparietal in dyscalculia². A few studies have already attempted to estimate the narrow sense heritability of brain activations for various tasks. They notably include digit and n-back working memory^{9,10}, visual math subtraction¹¹, and stimuli such as written words, faces and spoken language¹². However, these studies had relatively small sample sizes for reliably detecting heritability estimates ranging between 25 and 50%. The previous samples included 30 subjects (10 triplets of male monozygotic (MZ) twins with one additional brother)⁹, 64 subjects (19 MZ and 13 dizygotic (DZ) pairs)^{11,12} or 319 subjects (75 MZ and 66 DZ pairs, 37 unpaired)¹⁰. In addition to including a larger sample size, the HCP data were processed using state of the art methods, providing 2-mm isotropic resolution and finer inter-individual registration. In particular, the so-called grayordinate activations are computed on the surface of the cortex¹³ for each individual, and inter-subject fMRI alignment is performed using areal-feature-based registration¹⁴. The grayordinate approach refers to fMRI analyses performed on the cortical surface, as opposed to a volume-based approach. The same idea was applied to build the multimodal parcellation of the human cerebral

cortex¹⁵ onto which we decomposed our heritability analyses, enabling us to map the genetic influence on fMRI activations on a very fine scale.

Furthermore, it is known that neural activation endophenotypes from MRI may reflect not only impairments in language but also differences in cognitive scores. For example, weaker left-lateralizations have been reported for some developmental language disorders¹⁶, and in normal populations, fMRI activations during simple tasks correlate with various cognitive scores. Notably, single digit calculation fMRI activations are predictive of high school math scores¹⁷, and the fronto-parietal functional connectivity in children performing a task that required them to match Arabic numbers to an array of dots correlated with their score on a standardized math test¹⁸. Using HCP data and in line with these approaches, we show in this paper how variations in language related fMRI activations correlate with cognitive abilities assessed by the median reaction time (RT), average accuracy and difficulty level during the HCP language and math tasks. Remarkably, recent studies have shown that, beyond the age-related heritability of general cognitive ability^{19,20} and of various indicators of academic performance²¹, these scores are highly pleiotropic^{19,22} [pleiotropy occurs when one gene regulates one or more phenotypic traits].

This raises the question of the potential pleiotropy between neural activations and cognitive abilities. Thus, as a second contribution, we studied the shared genetic variance of fMRI activations and cognitive performance scores measured during the MRI session or behavioral scores acquired independently from the task. We studied behavioral variables measured by the HCP using standardized tests from the National Institute of Health (NIH): fluid intelligence, working memory, and language assessments such as vocabulary comprehension and oral reading decoding. Details of these variables and their heritability estimates can be found in Table S1, and how well they correlate phenotypically and genetically with the behavioral scores measured during the task is reported in Table S2.

Recent genome wide association studies have unveiled new loci and genes influencing human cognitive performance (e.g. human intelligence²³, general cognitive function^{24,25} and educational attainment²⁵) and possibly intelligence as a construct in differential psychology¹⁹. However, for these human-specific characteristics, little is known about the underlying integration mechanism of molecular functions or the brain areas where they are most influential. The shared genetic etiology investigated in this work provides new perspectives to decipher the basis of cognitive abilities such as language in humans. This study had two major

aims: (1) to estimate the heritability of fMRI activations during story comprehension and math tasks; and (2) to determine the shared genetic etiology between these activations and cognitive performance.

Results

Task fMRI Activations in MATH and STORY tasks

Figure 1 shows the activations for MATH (vs the intercept of the general linear model (GLM) being considered as baseline), STORY and the contrast STORY - MATH. The intercept reflects the mean of the residual BOLD time series after removing variance explained by all other regressors. Both tasks show clear activations in the planum temporale and Heschl's gyrus area, reflecting the fact that the stimuli were presented in the auditory modality. The MATH task, in which participants were requested to perform addition and subtraction, activates areas traditionally implicated in mathematical calculations, that is, the intraparietal sulcus, the middle frontal and the inferior temporal regions^{26,27}. The story listening task activates the language understanding network, encompassing bilateral temporal regions and left frontal regions^{28,29}. As expected, the group activations for the STORY task are more left lateralized, notably in the left posterior superior temporal and inferior frontal regions, which correspond to Wernicke's and Broca's areas, respectively. Moreover, regions implicated in inhibition networks are also activated by these tasks^{30,31}, notably the middle frontal gyrus in the math task and the medial prefrontal cortex, implicated in motivation and execution, and above the anterior cingulate cortex, controlling selective attention^{32,33}. In addition, both tasks activate complementary networks; in particular, the math task deactivates the semantic and episodic memory processes, known as the default mode network, which is also active in resting or passive states³⁴. This last remark makes the STORY-MATH contrast particularly relevant for studying the genetic influence on activation specifically elicited by math and story tasks.

Univariate Genetic Analyses

We performed a cortex-wise heritability analysis on the median activation (β -value) in the 360 areals of the HCP multi-modal parcellation. After stringent Bonferroni correction ($p < 0.05/360 \approx 1.4 \cdot 10^{-4}$), we found 54 regions whose activations during the MATH task are heritable and 46 regions for the STORY task. These results are summarized in **Figure 2** and heritability estimates are included in Tables S3-S6. The details and names of the areals can be found in the Supplementary Information of the paper describing the multimodal parcellation of the human cerebral cortex¹⁵. For the MATH (resp. STORY) task, the heritability estimates range from 0.23 to 0.45, with the maximum in the left "Area PGp" corresponding to the

angular gyrus (resp. 0.22 to 0.55, with the maximum in the left “PeriSylvian Language Area”). In addition, we performed heritability analysis using the median z-stat value in each areal instead of the median parameter estimate (β -value) and obtained similar results (Fig S1).

The univariate genetic analysis of the activations associated with verbal math emphasizes mainly areals spanning the math network, including the intraparietal sulcus, middle frontal, inferior temporal and angular gyri. The analysis of activations associated with story comprehension distinctively underlines regions of the language network as bilaterally heritable. Among these regions are the superior temporal sulcus dorsal and ventral parts, Brodmann area (BA) 47 in Broca’s area, and the middle frontal gyrus at the junction with the precentral sulcus. Interestingly, the heritability networks of the MATH and STORY tasks overlap very little except in the auditory cortex, around the planum temporale, in the frontal cortex (BA 8), and in the inferior temporal region.

Table 1 presents the heritability estimates of the behavioral scores gathered during the MRI scans. The global accuracy on the HCP language tasks, averaging the scores in the MATH and STORY tasks, was significantly heritable, with $h^2 = 0.34$, close to the traditionally high heritability estimate of cognitive performance^{35,36}. The heritability estimates for the median reaction time (RT) were approximately 0.2. Furthermore, RT Story and RT Math were significantly correlated (phenotypic correlation: $\rho_p = 0.34$, genetic correlation: $\rho_g = 0.45$, Table S2). Regarding the accuracy and average difficulty level of the HCP language task, we observed that the MATH task variables have higher heritability estimates than those of the STORY task. This result might indicate a higher genetic influence on performance during simple arithmetic tasks than during language comprehension. However, this result needs to be considered in light of the different distribution patterns of MATH and STORY accuracies. The STORY accuracy reported by the HCP displays discrete values and might not be sufficiently informative (Fig S2). Table S2 underlines a significant correlation between math and story accuracies ($\rho_p = 0.15$, $\rho_g = 0.27$). The discrete distribution of STORY accuracy likely occurs because each story lasted approximately 20-30s, few story questions were presented to the subjects, and most subjects tended to choose the correct answer in the two-alternative forced-choice question.

Bivariate Genetic Analyses

We performed bivariate genetic analyses to quantify the shared genetic influence between intellectual performance, represented by the behavioral measures, and the neural activation in

each areal. The genetic correlation estimates are usually subject to substantial sampling errors and therefore inaccurate. The large sample size of the HCP offers the opportunity to reduce the standard errors. The distribution of STORY accuracy values is concentrated on a small number of values (Fig S2), thus, we chose to use the average of the STORY and MATH accuracies as the behavioral score to characterize the individual performance. Thus, the results presented here concern the relationship between this average score and the activations or deactivations revealed by the contrast STORY-MATH (activation map shown in **Figure 1c**). As a first step of our analysis and to filter out the areals for which the neural activation was not significantly correlated with the behavioral score, we computed the phenotypic correlation between these two variables in each areal of the HCP multi-modal parcellation. **Figure 3 (a, b)** summarizes the phenotypic correlation and associated p-values, for areals significant after Bonferroni correction ($p < 0.05/360$). The language network is clearly encompassed along the left superior temporal sulcus (STS) and Broca's area, as well as in the anterior part of the right STS. The activations in the angular gyrus (area PGp), supporting the manipulation of numbers in verbal form³⁷, were also significantly correlated with the behavioral scores.

Among the 360 areals of the HCP multimodal parcellation, 39 (resp. 38) were significantly phenotypically correlated and were kept for the bivariate analysis in the right (resp. left) hemisphere. The shared genetic variance estimates for these areals are presented **Figure 3 (c, d)** ($p < 0.05$ without correction), and detailed values can be found in Tables S7 and S8. With stringent Bonferroni correction the p-value threshold for ρ_g is $p < 0.05/(39+38) \approx 6.5 \cdot 10^{-4}$. Among the areals with significant shared genetic variance, we found the left anterior ventral insular area (AVI, $\rho_g = 0.61$) and the right angular gyrus (PGp, $\rho_g = -0.40$). Noticeably, in the left hemisphere areals, parts of the language network in the posterior STS had activations that shared significant genetic variance with language accuracy. These include the posterior ventral (STSV posterior, $\rho_g = 0.47$) and dorsal (STSD posterior, $\rho_g = 0.45$) parts of the STS, adjacent to the auditory 5 complex area (A5, $\rho_g = 0.54$), the perisylvian language area (PSL, $\rho_g = 0.47$) and the temporo-intraparietal junction (PGi, $\rho_g = 0.54$). On the left hemisphere internal face, we also found the superior frontal language area (SFL, $\rho_g = 0.61$) and, adjacent to this areal the Brodmann 8 decomposed into medial (8Bm, $\rho_g = 0.75$) and lateral (8Bl, $\rho_g = 0.73$) parts. Additionally, we noted two right hemisphere regions implicated in language processing and significantly genetically correlated with the fMRI task average score: the temporal pole (area TG dorsal, TGd, $\rho_g = 0.65$) and the lateral part of Brodmann area 47 (47l, $\rho_g = 0.51$). The

latter is adjacent to Brodmann areas 44 and 45 in the inferior frontal, which are connected through the arcuate fasciculus with the language temporal regions.

Additionally, we extended our analysis to behavioral variables measured by the HCP following a standardized NIH protocol. Among these, we selected the variables that are most likely to reflect cognitive performance. Then, we estimated their heritability (Table S1) and correlations with the behavioral scores measured during the fMRI task (Table S2). In this set of variables, fluid intelligence (heritability: $h^2=0.43$, correlations with language accuracy: $\rho_p = 0.36$, $\rho_g = 0.61$), working memory ($h^2=0.52$, $\rho_p = 0.34$, $\rho_g = 0.50$), vocabulary comprehension ($h^2=0.64$, $\rho_p = 0.40$, $\rho_g = 0.57$) and oral reading decoding ($h^2=0.67$, $\rho_p = 0.46$, $\rho_g = 0.67$) were the ones with the highest heritability estimates and correlations with the average accuracy of the two fMRI tasks. Thus, we performed a bivariate genetic analysis between the STORY-MATH activations (difference between β_{STORY} and β_{MATH}) and these four variables. Regardless of whether one considers the NIH scores or the ones directly related to the fMRI tasks, the study of shared genetic influence with the median activation yields approximately the same set of regions (**Figures 4, 5**). This observation reinforces our claim that these regions have common genetic roots with the parts of general cognitive performance accounted for by the four cognitive variables under scrutiny, namely fluid intelligence, working memory, vocabulary comprehension and reading decoding.

Discussion

In this paper, we have shown that brain activation pattern in the language and math networks are heritable. Additionally, we highlighted a particular set of regions along the superior temporal sulcus and in the inferior frontal whose activations share a common genetic basis with some aspects of general cognitive ability, assessed through fMRI task accuracy and behavioral scores.

We must emphasize that these results correspond to fMRI activations associated with verbal math and semantic comprehension tasks. Thus, regions not recruited by the tasks cannot be found to be significantly correlated with cognitive ability in our case, because activations in these regions are incoherent across individuals. Notably, the visual word form area, related to literacy, is not activated in our oral tasks because they did not require word reading.

Furthermore, combining data from various cohorts is unfeasible, because neural activations from different tasks are not comparable when estimating inter-individual variance. This highlights the necessity of utilizing large cohorts with standardized fMRI protocols to perform such genetic analyses. To our knowledge, this is the first study to address the heritability of fMRI activation cortex-wise on a multimodal parcellation of the human cerebral cortex. Our results confirm the genetic influence on the formation of neural circuits implicated in language³⁸ and math¹¹. Using the HCP fine scale parcellation¹⁵ allowed us, for instance, to distinguish the genetic effects on the temporo-parietal junction implicated in language²⁸ (area PGI) and on the adjacent angular gyrus (area PGp), which is particularly involved in the manipulation of numbers in verbal form³⁷. Indeed, these two areas, part of BA 39, present different cytoarchitectonic properties, such as a slightly broader layer II for PGI³⁹, which might explain their involvement in different tasks. In a previous work, Pinel and Dehaene also found the left angular gyrus and the posterior superior parietal lobule bilaterally to be heritable¹¹. Adding to these observations, our results underline a left hemisphere intraparietal specificity, with more heritable areals and slightly higher heritability compared to the right for the MATH contrast. This finding is consistent with results reported by Vogel and colleagues demonstrating a correlation of activations in the left intraparietal sulcus modulated by age, which was not observed in the right intraparietal⁴⁰. Heritability represents the proportion of observed inter-individual phenotypic variance that is explained by genetics. Thus, it might be that inter-individual variance is not sufficiently pronounced in the right hemisphere, whereas activations have evolved over one's lifetime in the left hemisphere. Overall, the heritability

maps for the STORY and MATH tasks pinpoint regions known to be disrupted in neurodevelopmental disorders. For instance, the inferior frontal area and the temporo-parietal junction activations are impaired in developmental dyslexia^{1,41}, and the intraparietal region activations are less modulated by the numerical distance between two numbers being compared in developmental dyscalculia^{2,42,43}. Highlighted areas might provide new insights into brain regions where normal gene expression might be disrupted, leading to brain dysfunction and neurodevelopmental disorders. Frequently replicated genes associated with neurobehavioral disorders, such as developmental dyslexia or SLI, likely play such a role in structural brain maturation by interfering with neuronal migration and neurite growth⁴.

Several studies have already described some phenotypic correlations between cognitive abilities and neural activations in language^{44–46} and math^{17,18,47,48}. Our study replicates these observations, notably the correlation with language processing regions, including Broca's area and the posterior superior temporal gyrus⁴⁴. Moreover, we estimated the genetic proportion in these phenotypic correlations. Hence, we demonstrated a shared genetic etiology between brain activations and cognitive performance, assessed in our study by the following tests: fluid intelligence, working memory, vocabulary comprehension and reading decoding. Interestingly, in the right hemisphere, mainly the anterior STS was found to be genetically correlated with the language task accuracy. This result seems consistent with the hypothesized role of the right anterior STS in the processing of prosody or figurative language, likely involved in the Aesop's fable metaphors presented to the subjects⁴⁹.

The observed genetic correlations shed light on the genetic links between cognitive performance and activation level in cognitive task-related fMRI. These links might be related to the development and maturation of myelin, enhancing brain connectivity. In children with difficulties processing syntactically complex sentences, arcuate fasciculus maturation was incomplete compared to adults^{50,51}. Thus, we could look for additive genetic effects implicated in the various levels of fiber tract maturation, which improves brain connectivity and efficiency. Indeed, Skeide and colleagues reported an example of such a genetic risk variant for dyslexia. They showed that this variant is related to the functional connectivity of left fronto-temporal phonological processing areas during the resting state⁵². Similarly, children with higher arithmetic scores present a more mature response modulation in their left intraparietal lobe⁴⁷. Our study suggests that a proportion of the observed inter-individual variance in cognitive performance partly results from the same additive genetic effects as those contributing to brain activation variance. The moderate shared genetic basis suggests

that a crucial interaction occurs between the environment and gene networks to enable the brain to develop to its full potential.

Recently, with the emergence of large cohorts, such as UKBiobank, new loci and genes influencing human cognitive ability have been discovered²³⁻²⁵. However, little is known about how these genes contribute to this human-specific trait. Our study pinpoints brain regions where activations genetically correlate with global cognition scores. These regions might help elucidate the mechanism in which these genes are implicated. When the HCP genotyping data are released, a polygenic score of these newly discovered variants could be used to determine the explained proportion of the neural activation variance in these regions.

Material and Methods

Subjects

This study utilized the dataset of the Human Connectome Project (HCP). The HCP scans and data were released in April 2017 (humanconnectome.org). The details of the release are available in the HCP reference manual. In this project, 1046 subjects aged between 22 and 37 years old ($\mu \pm \sigma = 28.8 \pm 3.7$ years) completed the fMRI language task in the HCP S1200 release. To avoid population stratification, we only included the 785 Caucasian individuals (372/413 M/F) that were classified as race="White" by the HCP; among these, the ethnicity of 69 was "Hispanic/Latino". This subgroup of the HCP contains 178 twin pairs (117 monozygotic twins (MZ) with 103 siblings and 61 dizygotic twins (DZ) with 61 siblings and 1 half sibling), 203 siblings, 1 half sibling and 60 unpaired individuals. The unpaired individuals did not contribute to the genetic parameter estimation but allowed a more accurate estimation of mean and variance effects. Subjects were chosen by the HCP consortium to represent healthy adults beyond the age of major neurodevelopmental changes and before the onset of neurodegenerative changes⁵³. They underwent a battery of tests to determine if they met the inclusion/exclusion criteria of the HCP⁵³. All subjects provided written informed consent on forms approved by the Institutional Review Board of Washington University. All of the following methods were carried out in accordance with relevant guidelines and regulations. All experimental MRI protocols were approved by the Institutional Review Board of Washington University.

Data availability statement: HCP (<https://db.humanconnectome.org>) is a publicly available dataset. Investigators need to apply to be granted access to restricted data.

Statistical power analysis

The statistical power of an extended pedigree study should not be confused with that of a cohort of unrelated subjects. In the latter case, the software often used is GCTA (genome wide complex trait analysis)⁵⁴, which enables one to easily compute the statistical power⁵⁵. In an extended pedigree, the statistical power computation is a hard problem because non-independence among relatives must be taken into account⁵⁶. To the best of our knowledge, there is no consensus method to compute the *a priori* statistical power for studies on imaging genetics in the case of a large pedigree cohort. For the sake of comparison, we propose to compare our study with previously published studies. Previous heritability studies that used approximately 800 subjects from the HCP, roughly as much as in our sample, obtained standard errors between 0.06 and 0.08 for phenotypes with heritability between 0.20 and 0.45^{57,58}. Using the GCTA Power Calculator⁵⁵ with the default value for the variance of the single nucleotide polymorphism (SNP) derived genetic relationship matrix of 10^{-5} , this range of standard errors and heritability estimates requires a sample size of 6000-7000 unrelated individuals. To allow

comparison with other pedigrees studies, we computed a relatedness summary parameter (RSP) using the pedRSP R package⁵⁹. We determined that our sample of 785 individuals is equivalent to an effective number of 1,922 pairs of siblings (RSP of our sample: $V_\lambda = 0.649$, $V_q = 0.671$, $\gamma = 0.73$; see⁵⁹ for details of the parameter's meaning).

Additionally, we want to emphasize that modeling non-additive genetic effects⁶⁰ is not possible when genetic data are not available. In a pedigree study such as the one currently available from the HCP, the only information available is the relationship between individuals.

Image acquisition and processing

MR images were acquired using a 3T Connectome Scanner, adapted from Siemens Skyra, housed at Washington University in St Louis, using a 32-channel head coil. T1-weighted images with 256 slices per slab were acquired with the three-dimensional magnetization-prepared rapid gradient echo (3D-MPRAGE) sequence: TR=2400 ms, TE=2.14 ms, TI=1000 ms, flip angle=8°, FOV=224×224 mm, and resolution 0.7 mm isotropic. T2-weighted images, 256 slices per slab, were acquired with a 3D-T2SPACE sequence: TR=3200 ms, TE=565 ms, variable flip angle, FOV=224×224 mm, and resolution 0.7 mm isotropic. fMRI data acquisition parameters were as follows: TR=720 ms, TE=33.1 ms, flip angle=52 deg, BW=2290 Hz/Px, in-plane FOV=208×180 mm, 72 slices, and 2.0 mm isotropic voxels, with a multi-band acceleration factor of 8. Two runs of each task were acquired, one with right-to-left and the other with left-to-right phase encoding 2; each run interleaved 4 blocks of a story task with 4 blocks of a math task. The lengths of the blocks varied (average of approximately 30 seconds), but the task was designed so that the math task blocks matched the length of the story task blocks, with some additional math trials at the end of the task to complete the 3:57 (min:sec) run.

The details of the HCP data analysis pipelines are described elsewhere^{13,61}. Briefly, they are primarily built using tools from FSL⁶² and Freesurfer⁶³. The HCP fMRIVolume pipeline generates “minimally preprocessed” 4D time series that include gradient unwarping, motion correction, fieldmap-based EPI distortion correction, brain-boundary-based registration of EPI to structural T1-weighted scans, non-linear (FNIRT) registration into MNI152 space, and grand-mean intensity normalization⁶¹. For the S500 release, two smoothing approaches were chosen by the HCP: volume-based smoothing or smoothing constrained to the cortical surface and subcortical gray-matter parcels. For the former, standard FSL tools can be applied for analysis, while for the latter, the HCP adapted these tools to this the ‘grayordinate’ approach^{13,61}. The grayordinate approach refers to fMRI analyses performed on the cortical surfaces, as opposed to a volume-based approach. This is more accurate spatially because activation occurs in gray, not white, matter. Unconstrained volume-based smoothing causes blurring effects by mixing signals from cortex regions adjacent in volume but not on the surface. For these reasons, our study analyses were carried out on the surface of the cortex.

The HCP fMRISurface pipeline brings the time series from the volume into the CIFTI grayordinate standard space. This is accomplished by mapping the voxels within the cortical gray matter ribbon onto the native cortical surface, transforming them according to the surface registration onto the 32k Conte69 mesh, and mapping the set of subcortical gray matter voxels from each subcortical parcel in each individual to a standard set of voxels in each atlas parcel. The result is a standard set of grayordinates in every subject (i.e., the same number in each subject, with spatial correspondence) with 2mm average surface vertex and subcortical volume voxel spacing. These data are smoothed with surface and parcel constrained smoothing of 2mm FWHM (full width half maximum) to regularize the mapping process⁶¹.

The language task

The HCP language task was developed by Binder and colleagues³⁴. In the story blocks, participants were presented with brief auditory stories adapted from Aesop's fables, followed by a 2-alternative forced-choice question to check the participants' understanding of the story topic. The example provided in the original paper is "*For example, after a story about an eagle that saves a man who had done him a favor, participants were asked, "Was that about revenge or reciprocity?"*". In the math blocks, participants were also presented auditory series of addition and subtraction (e.g., "fourteen plus twelve"), followed by "equals" and then two choices (e.g., "twenty-nine or twenty-six"). To ensure similar level of difficulty across participants, math trials automatically adapted to the participants responses. As shown by Binder and colleagues³⁴, the story and math trials were well matched in terms of duration, auditory and phonological input, and attention demand. Furthermore, they were likely to elicit distinct brain activation – on the one hand, anterior temporal lobes classically involved in semantic processing, and parietal cortex on the other hand, classically involved in numerical processing, thus spanning a broad set of regions involved in conceptual semantic processing.

HCP task fMRI analysis

The analysis of fMRI data was carried out by the HCP consortium and we describe briefly their pipeline¹³. The Story predictor covered the variable duration of a short story, question, and response period (~30 s). The Math predictor covered the duration of a set of math questions designed to roughly match the duration of the story blocks. The grayordinate data for individual task runs were processed in a level 1 analysis. Activity estimates were computed for the preprocessed functional time series from each run using a general linear model (GLM) implemented in FSL's FILM (FMRIB's Improved Linear Model with autocorrelation correction)⁶⁴. Predictors were convolved with a double gamma "canonical" hemodynamic response function⁶⁵ to generate the main model regressors. The two runs for each task and subject were then combined in a level 2 fixed-effects analysis¹³, which we used as our phenotype. Fixed-effects analyses were conducted using FEAT (fMRI Expert Analysis Tool) to estimate the average effects across runs within-subjects, and then mixed-effects analyses treating

subjects as random effects were conducted using FLAME (FMRIB's Local Analysis of Mixed Effects) to estimate the average effects of interest for the group third-level analysis.

Phenotype definitions

To define our phenotypes, we consider separately the regression analyses on STORY and MATH tasks, and the contrast STORY-MATH. We used the beta values (pe1.dtseries.nii files) of the results of the level 2 analysis, which essentially average the level 1, i.e., the individual, runs. The contrasts were defined by the HCP in level 1 and averaged for level 2: thus, the grayordinate values of the beta and contrast values (cope1.dtseries.nii) are identical in this case, as they did not define any “new” contrasts specifically at level 2. Therefore, we could have used the cope1.dtseries.nii files with no difference in results. We used the MSMAll registered the functional analysis results from HCP and the HCP multimodal parcellation¹⁵. We analyzed each of the 180 areals separately. We computed the median beta values in each areal for both hemispheres. These phenotypes constitute our proxy to estimate the activation in each part of the brain.

Moreover, we also included in our phenotypes the accuracy, reaction time and average difficulty level for the MATH and STORY tasks. We called these “behavioral scores”, as opposed to the grayordinate activation phenotypes previously defined.

Univariate analysis of additive genetic variance

The variance components method, as implemented in the Sequential Oligogenic Linkage Analysis Routines (SOLAR) software package⁶⁶, was used for the heritability estimations of the phenotypes under analysis, such as the median activation in each areal¹⁵. The SOLAR algorithms use maximum variance decomposition methods derived from the strategy developed by Amos⁶⁷. The covariance matrix Ω for a pedigree of individuals is given by:

$\Omega = 2 \cdot \Phi \cdot \sigma_g^2 + I \cdot \sigma_e^2$, where σ_g^2 is the genetic variance due to the additive genetic factors, Φ is the kinship matrix representing the pair-wise kinship coefficients among all individuals, σ_e^2 is the variance due to individual-specific environmental effects, and I is the identity matrix.

Narrow sense heritability is defined as the fraction of the phenotype variance σ_p^2 attributable to additive genetic factors: $h^2 = \sigma_g^2 / \sigma_p^2$.

The significance of the heritability is tested by comparing the likelihood of the model in which σ_g^2 is constrained to zero with that of the model in which σ_g^2 is estimated. Before testing for the significance of heritability, phenotype values for each individual within the HCP cohort were adjusted for the following covariates: sex, age, age², age·sex interaction, age²·sex interaction, ethnicity (Hispanic or not) and education level. We used the number of years of education as a proxy for the education level to account for environmental differences in family socioeconomic status. This is a conservative

approach because the number of years of education was shown to be associated not only with the family socioeconomic status (7%) but also with the general cognitive ability (3.5%)⁶⁸. Thus, it likely has shared environmental ground with the former and shared genetic origin with the latter. HCP data do not contain the information that would disentangle this issue. Following this last remark, one should note that the heritability estimates and shared genetic variances, described in the next section, were underestimated.

Bivariate genetic analyses

To assess the relationship between math dexterity/language comprehension and activation in brain areas, we computed the Pearson correlation between the median activation in each of the 180 areals of both hemispheres, and the behavioral scores.

Furthermore, we assessed the degree of shared genetic variance in the areals for which activation was significantly correlated with the behavioral scores; we performed a genetic correlation analysis using SOLAR, relying on the following model:

$$\rho_p = \sqrt{h_a^2} \sqrt{h_b^2} \cdot \rho_g + \sqrt{1 - h_a^2} \sqrt{1 - h_b^2} \cdot \rho_e$$
, where Pearson's phenotypic correlation ρ_p is decomposed into ρ_g and ρ_e . ρ_g is the proportion of variability due to shared genetic effects and ρ_e that due to the environment, while h_a^2 and h_b^2 correspond to the previously defined narrow sense heritability for phenotypes a and b , respectively. In our case, one corresponds to the heritability of fMRI activation in one areal, while the second is the heritability of one of our behavioral scores.

Acknowledgments

Data were provided by the Human Connectome Project, WU-Minn Consortium (Principal Investigators: David Van Essen and Kamil Ugurbil; 1U54MH091657) funded by the 16 NIH Institutes and Centers that support the NIH Blueprint for Neuroscience Research; and by the McDonnell Center for Systems Neuroscience at Washington University.

Author Contributions Statement

Y.L.G and V.F designed the analysis and wrote the manuscripts. Y.L.G prepared the figures and tables. All authors analyzed the results and reviewed the manuscripts.

Additional Information

Competing financial and non-financial interests

All the co-authors have no conflict of interest to declare.

References

1. Peterson, R. L. & Pennington, B. F. Developmental dyslexia. *Lancet* **379**, 1997–2007 (2012).
2. Price, G. & Ansari, D. Dyscalculia: Characteristics, Causes, and Treatments. *Numeracy* **6**, (2013).
3. Docherty, S. J. *et al.* A genome-wide association study identifies multiple loci associated with mathematics ability and disability. *Genes, Brain Behav.* **9**, 234–247 (2010).
4. Mascheretti, S. *et al.* Neurogenetics of developmental dyslexia: from genes to behavior through brain neuroimaging and cognitive and sensorial mechanisms. *Transl. Psychiatry* **7**, e987 (2017).
5. Männel, C. *et al.* Working-memory endophenotype and dyslexia-associated genetic variant predict dyslexia phenotype. *Cortex* **71**, 291–305 (2015).
6. Friederici, A. D. The Neural Basis of Language Development and Its Impairment. *Neuron* **52**, 941–952 (2006).
7. van der Lely, H. K. J. & Pinker, S. The biological basis of language: Insight from developmental grammatical impairments. *Trends Cogn. Sci.* **18**, 586–595 (2014).
8. Leonard, L. B. in *Language competence across populations: Toward a definition of specific language impairment.* (2003).
9. Koten, J. W. *et al.* Genetic Contribution to Variation in Cognitive Function: An fMRI Study in Twins. *Science (80-.).* **323**, 1737–1740 (2009).
10. Blokland, G. a *et al.* Heritability of working memory brain activation. *J Neurosci* **31**, 10882–10890 (2011).
11. Pinel, P. & Dehaene, S. Genetic and environmental contributions to brain activation during calculation. *Neuroimage* **81**, 306–316 (2013).
12. Pinel, P. *et al.* Genetic and environmental influences on the visual word form and fusiform face areas. *Cereb. Cortex* **25**, 2478–2493 (2015).
13. Barch, D. M. *et al.* Function in the human connectome: Task-fMRI and individual differences in behavior. *Neuroimage* **80**, 169–189 (2013).
14. Robinson, E. C. *et al.* MSM: A new flexible framework for multimodal surface matching. *Neuroimage* **100**, 414–426 (2014).

15. Glasser, M. F. *et al.* A multi-modal parcellation of human cerebral cortex. *Nature* **536**, 171–8 (2016).
16. Bishop, D. V. M. Cerebral Asymmetry and Language Development: Cause, Correlate, or Consequence? *Science* (80-.). **340**, 1230531–1230531 (2013).
17. Price, G. R., Mazzocco, M. M. M. & Ansari, D. Why Mental Arithmetic Counts: Brain Activation during Single Digit Arithmetic Predicts High School Math Scores. *J. Neurosci.* **33**, 156–163 (2013).
18. Emerson, R. W. & Cantlon, J. F. Early math achievement and functional connectivity in the fronto-parietal network. *Dev. Cogn. Neurosci.* **2**, S139–S151 (2012).
19. Plomin, R. & Deary, I. J. Genetics and intelligence differences: five special findings. *Mol. Psychiatry* **20**, 98–108 (2015).
20. Davies, G. *et al.* Genome-wide association studies establish that human intelligence is highly heritable and polygenic. *Mol. Psychiatry* **16**, 996–1005 (2011).
21. Haworth, C. M. A., Dale, P. & Plomin, R. A Twin Study into the Genetic and Environmental Influences on Academic Performance in Science in nine-year-old Boys and Girls. *Int. J. Sci. Educ.* **30**, 1003–1025 (2008).
22. Rimfeld, K., Kovas, Y., Dale, P. S. & Plomin, R. Pleiotropy across academic subjects at the end of compulsory education. *Sci. Rep.* **5**, 11713 (2015).
23. Sniekers, S. *et al.* Genome-wide association meta-analysis of 78,308 individuals identifies new loci and genes influencing human intelligence. *Nat. Genet.* **49**, 1107–1112 (2017).
24. Trampush, J. W. *et al.* GWAS meta-analysis reveals novel loci and genetic correlates for general cognitive function: a report from the COGENT consortium. *Mol. Psychiatry* **22**, 336–345 (2017).
25. Davies, G. *et al.* Genome-wide association study of cognitive functions and educational attainment in UK Biobank (N= 112 151). *Mol. Psychiatry* **21**, 1–10 (2016).
26. Amalric, M. & Dehaene, S. Origins of the brain networks for advanced mathematics in expert mathematicians. *Proc. Natl. Acad. Sci.* **113**, 4909–4917 (2016).
27. Rickard, T. . *et al.* The calculating brain: an fMRI study. *Neuropsychologia* **38**, 325–335 (2000).
28. Pallier, C., Devauchelle, A.-D. & Dehaene, S. Cortical representation of the constituent

- structure of sentences. *Proc. Natl. Acad. Sci.* **108**, 2522–2527 (2011).
29. Skeide, M. A. & Friederici, A. D. The ontogeny of the cortical language network. *Nat. Rev. Neurosci.* **17**, 323–332 (2016).
 30. Simmonds, D. J., Pekar, J. J. & Mostofsky, S. H. Meta-analysis of Go/No-go tasks demonstrating that fMRI activation associated with response inhibition is task-dependent. *Neuropsychologia* **46**, 224–232 (2008).
 31. Erika-Florence, M., Leech, R. & Hampshire, a. A functional network perspective on response inhibition and attentional control. *Nat Commun* **5**, 4073 (2014).
 32. Leisman, G., Machado, C., Melillo, R. & Mualem, R. Intentionality and ‘free-will’ from a neurodevelopmental perspective. *Front. Integr. Neurosci.* **6**, 1–12 (2012).
 33. Etkin, A., Egner, T. & Kalisch, R. Emotional processing in anterior cingulate and medial prefrontal cortex. *Trends Cogn. Sci.* **15**, 85–93 (2011).
 34. Binder, J. R. *et al.* Mapping anterior temporal lobe language areas with fMRI: A multicenter normative study. *Neuroimage* **54**, 1465–1475 (2011).
 35. Shakeshaft, N. G. *et al.* Strong Genetic Influence on a UK Nationwide Test of Educational Achievement at the End of Compulsory Education at Age 16. *PLoS One* **8**, e80341 (2013).
 36. Krapohl, E. *et al.* The high heritability of educational achievement reflects many genetically influenced traits, not just intelligence. *Proc. Natl. Acad. Sci.* **111**, 15273–15278 (2014).
 37. Dehaene, S., Piazza, M., Pinel, P. & Cohen, L. THREE PARIETAL CIRCUITS FOR NUMBER PROCESSING. *Cogn. Neuropsychol.* **20**, 487–506 (2003).
 38. Vernes, S. C. & Fisher, S. E. in *Animal Models of Speech and Language Disorders* (ed. Helekar, S. A.) 13–40 (Springer New York, 2013).
 39. Caspers, S. *et al.* The human inferior parietal cortex: Cytoarchitectonic parcellation and interindividual variability. *Neuroimage* **33**, 430–448 (2006).
 40. Vogel, S. E., Goffin, C. & Ansari, D. Developmental specialization of the left parietal cortex for the semantic representation of Arabic numerals: An fMR-adaptation study. *Dev. Cogn. Neurosci.* **12**, 61–73 (2015).
 41. Eicher, J. D. & Gruen, J. R. Imaging-genetics in dyslexia: Connecting risk genetic variants to brain neuroimaging and ultimately to reading impairments. *Mol. Genet. Metab.* **110**, 201–212 (2013).

42. Price, G. R., Holloway, I., Räsänen, P., Vesterinen, M. & Ansari, D. Impaired parietal magnitude processing in developmental dyscalculia. *Curr. Biol.* **17**, 1042–1043 (2007).
43. Ashkenazi, S., Rosenberg-Lee, M., Tenison, C. & Menon, V. Weak task-related modulation and stimulus representations during arithmetic problem solving in children with developmental dyscalculia. *Dev. Cogn. Neurosci.* **2**, S152–S166 (2012).
44. Skeide, M. A., Brauer, J. & Friederici, A. D. Brain Functional and Structural Predictors of Language Performance. *Cereb. Cortex* **26**, 2127–2139 (2016).
45. van Ettinger-Veenstra, H. M. *et al.* Right-hemispheric brain activation correlates to language performance. *Neuroimage* **49**, 3481–3488 (2010).
46. Booth, J. R. *et al.* Relation between brain activation and lexical performance. *Hum. Brain Mapp.* **19**, 155–169 (2003).
47. Bugden, S., Price, G. R., McLean, D. A. & Ansari, D. The role of the left intraparietal sulcus in the relationship between symbolic number processing and children’s arithmetic competence. *Dev. Cogn. Neurosci.* **2**, 448–457 (2012).
48. Emerson, R. W. & Cantlon, J. F. Continuity and change in children’s longitudinal neural responses to numbers. *Dev. Sci.* **18**, 314–326 (2015).
49. Diaz, M. T. & Hogstrom, L. J. The influence of context on hemispheric recruitment during metaphor processing. *J. Cogn. Neurosci.* **23**, 3586–3597 (2011).
50. Friederici, A. D. & Gierhan, S. M. The language network. *Curr. Opin. Neurobiol.* **23**, 250–254 (2013).
51. Brauer, J., Anwander, A. & Friederici, A. D. Neuroanatomical prerequisites for language functions in the maturing brain. *Cereb. Cortex* **21**, 459–466 (2011).
52. Skeide, M. A. *et al.* Genetic dyslexia risk variant is related to neural connectivity patterns underlying phonological awareness in children. *Neuroimage* **118**, 414–421 (2015).
53. Van Essen, D. C. *et al.* The Human Connectome Project: A data acquisition perspective. *Neuroimage* **62**, 2222–2231 (2012).
54. Yang, J., Lee, S. H., Goddard, M. E. & Visscher, P. M. GCTA: A Tool for Genome-wide Complex Trait Analysis. *Am. J. Hum. Genet.* **88**, 76–82 (2011).
55. Visscher, P. M. *et al.* Statistical power to detect genetic (co) variance of complex traits using SNP data in unrelated samples. *PLoS Genet.* **10**, e1004269 (2014).

56. Blangero, J. *et al.* A Kernel of Truth. *Statistical Advances in Polygenic Variance Component Models for Complex Human Pedigrees. Advances in Genetics* **81**, (2013).
57. Le Guen, Y. *et al.* Genetic Influence on the Sulcal Pits: On the Origin of the First Cortical Folds. *Cereb. Cortex* 1–12 (2017). doi:10.1093/cercor/bhx098
58. Hodgson, K. *et al.* Shared Genetic Factors Influence Head Motion During MRI and Body Mass Index. *Cereb. Cortex* 1–8 (2016). doi:10.1093/cercor/bhw321
59. Raffa, J. D. & Thompson, E. A. Power and Effective Study Size in Heritability Studies. *Stat. Biosci.* **8**, 264–283 (2016).
60. Ge, T., Feng, J., Hibar, D. P., Thompson, P. M. & Nichols, T. E. Increasing power for voxel-wise genome-wide association studies: The random field theory, least square kernel machines and fast permutation procedures. *Neuroimage* **63**, 858–873 (2012).
61. Glasser, M. F. *et al.* The minimal preprocessing pipelines for the Human Connectome Project. *Neuroimage* **80**, 105–124 (2013).
62. Jenkinson, M., Beckmann, C. F., Behrens, T. E. J., Woolrich, M. W. & Smith, S. M. Fsl. *Neuroimage* **62**, 782–790 (2012).
63. Fischl, B. FreeSurfer. *Neuroimage* **62**, 774–781 (2012).
64. Woolrich, M. W., Ripley, B. D., Brady, M. & Smith, S. M. Temporal Autocorrelation in Univariate Linear Modeling of fMRI Data. *Neuroimage* **14**, 1370–1386 (2001).
65. Glover, G. H. Deconvolution of Impulse Response in Event-Related BOLD fMRI. *Neuroimage* **9**, 416–429 (1999).
66. Almasy, L. & Blangero, J. Multipoint Quantitative-Trait Linkage Analysis in General Pedigrees. *Am. J. Hum. Genet.* **62**, 1198–211 (1998).
67. Amos, C. I. Robust variance-components approach for assessing genetic linkage in pedigrees. *Am. J. Hum. Genet.* **54**, 535–543 (1994).
68. Selzam, S. *et al.* Predicting educational achievement from DNA. *Mol. Psychiatry* **22**, 267–272 (2017).

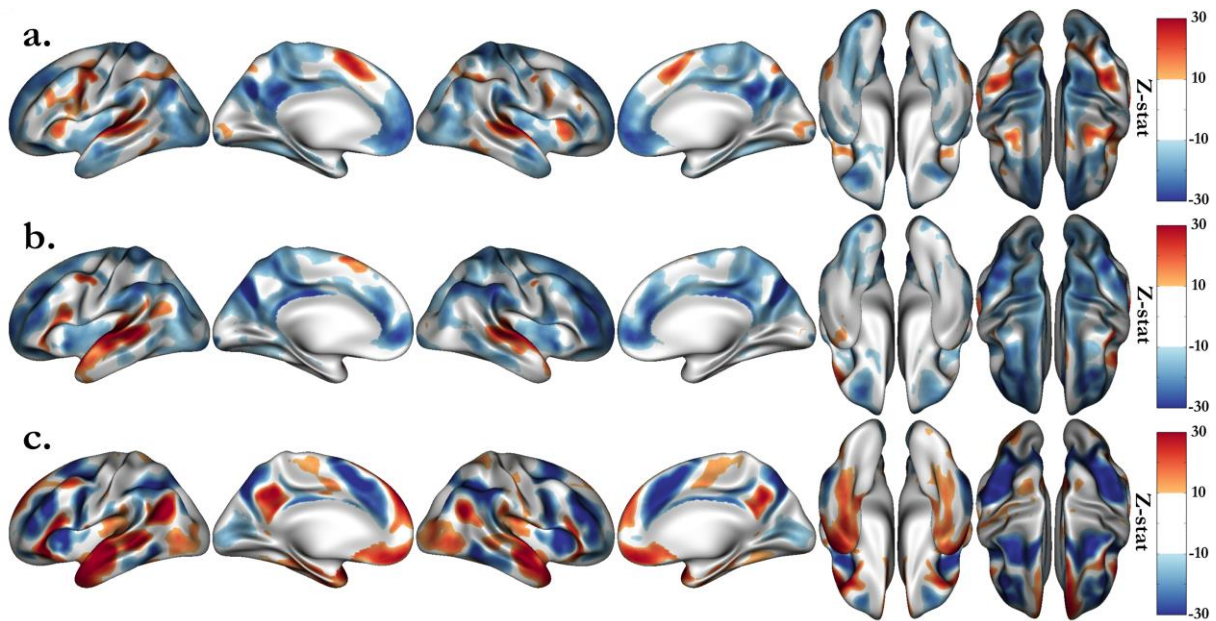


Figure 1. Group average activations for the HCP language tasks, including MATH (a.) and STORY (b.) blocks, and the STORY-MATH contrast (c.). Group maps are shown with a lower threshold of $z = \pm 10$ and saturation from $z = 30$ to introduce the main areas activated by the tasks. Due to the large number of subjects, the associated p-values are significant; we arbitrarily set the thresholds to emphasize the regions that are known to be recruited by these tasks.

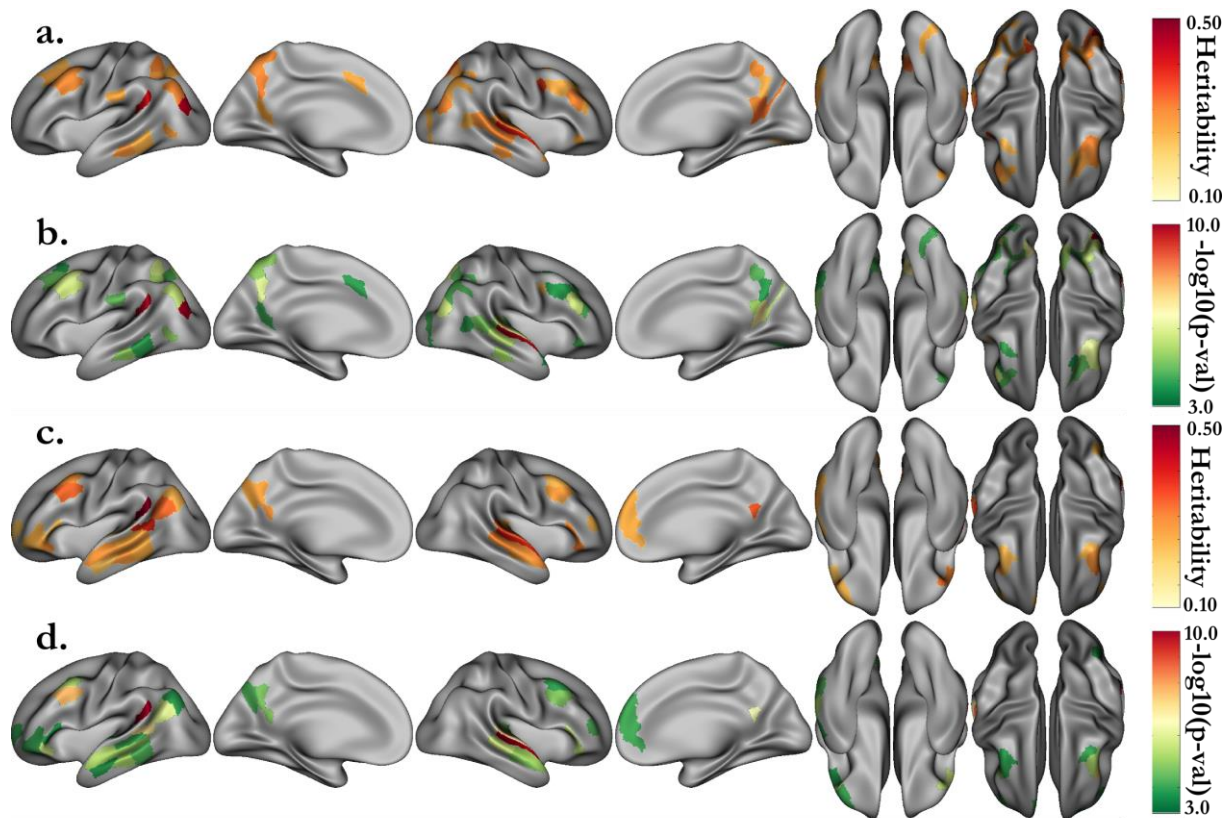


Figure 2. Heritability estimates for the activations of the MATH (a.) and STORY (c.) tasks, and their associated p-values (respectively, b. and d.). Only the estimates significant after correction ($p < 0.05/360$, with 180 areals in each hemisphere) are displayed. Activations correspond to the median parameter estimate (β) in each areal of the HCP multimodal parcellation.

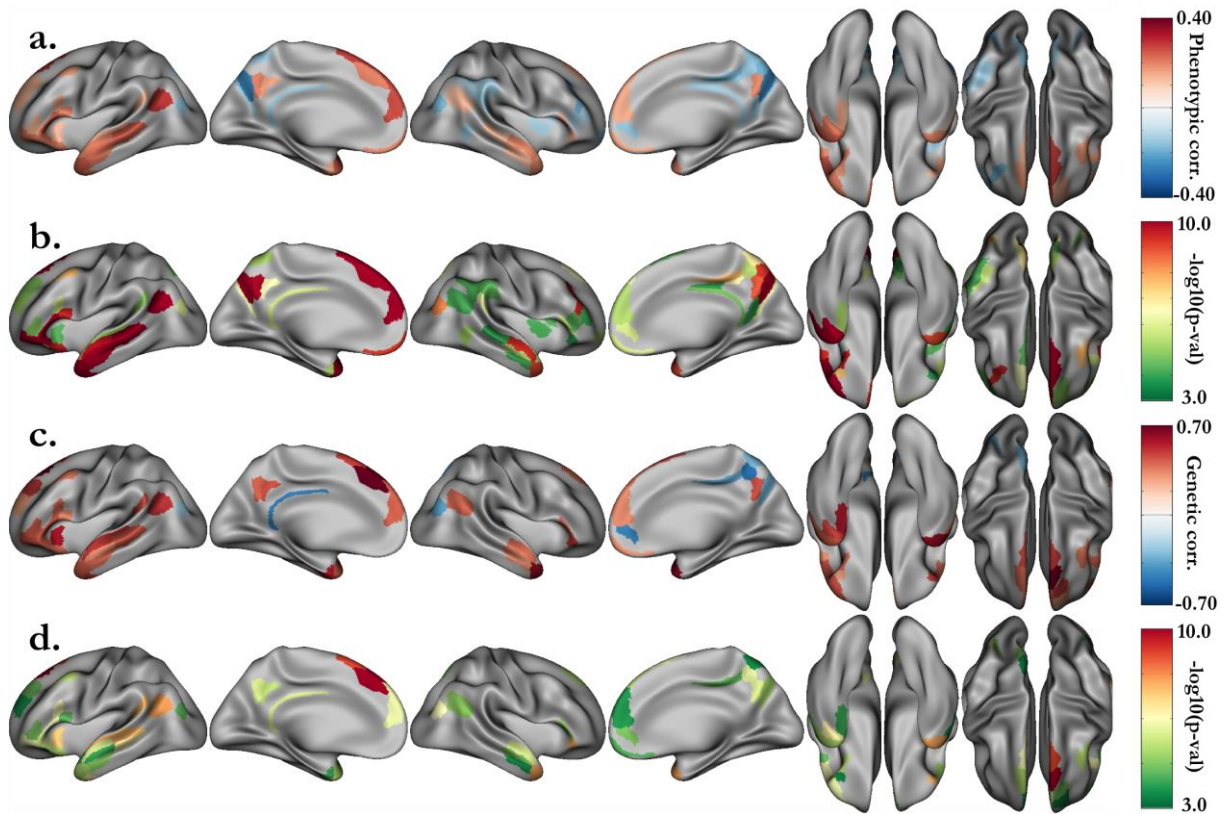


Figure 3. Bivariate genetic analysis results between HCP LANGUAGE task accuracy and activation for the STORY-MATH contrast in each areal. a. After strict Bonferroni correction ($p < 0.05/360$), significant phenotypic correlations between the language task accuracy (average of story and math accuracies) and the median activation of the contrast STORY-MATH in each areal, **b.** with their associated p-values. **c.** Proportion of variability due to shared genetic effects with **d.** their associated uncorrected p-values < 0.05 .

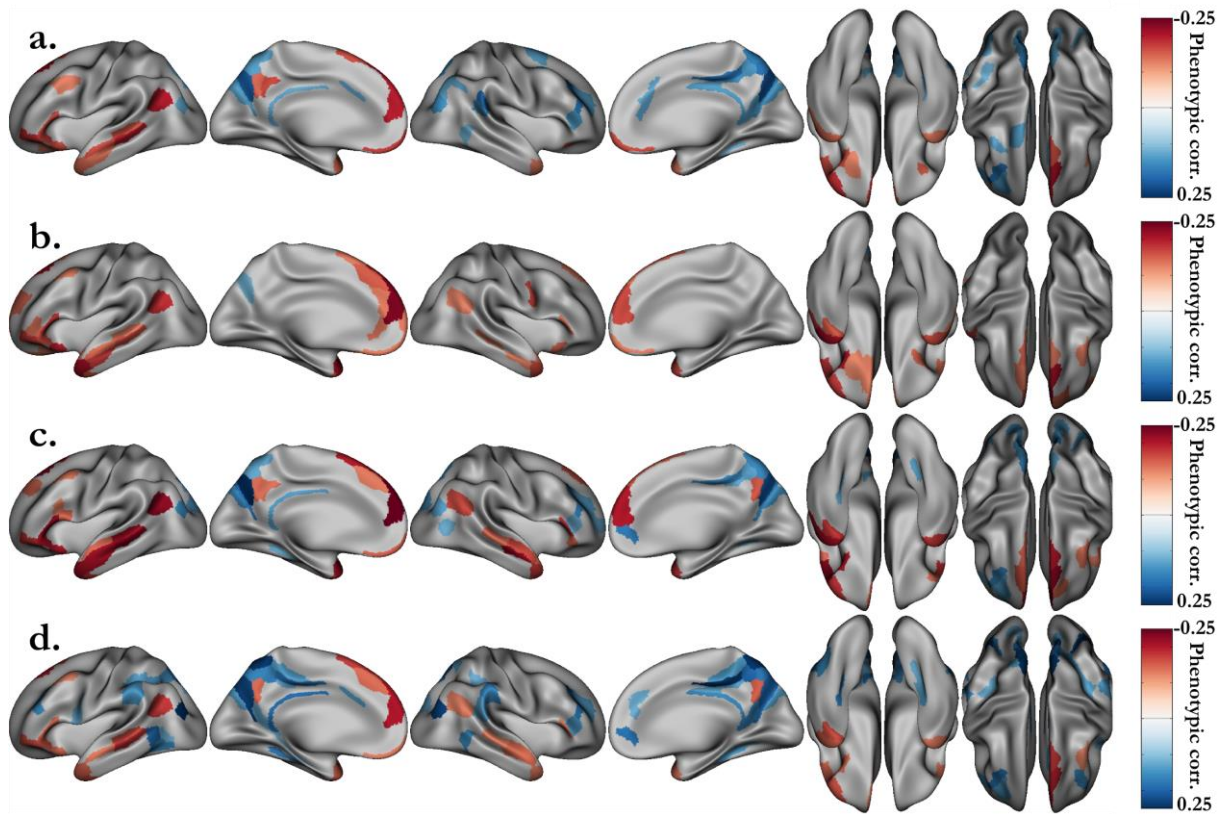


Figure 4. Significant phenotypic correlations between the grayordinate activations of the STORY-MATH contrast and the NIH behavioral scores. a. Fluid intelligence (PMAT24_A_CR). **b.** Working memory (ListSort). **c.** Vocabulary comprehension (PicVocab). **c.** Reading decoding (ReadEng). Associated p-values ($p < 0.05/360$, Bonferroni correction) can be found Fig S3.

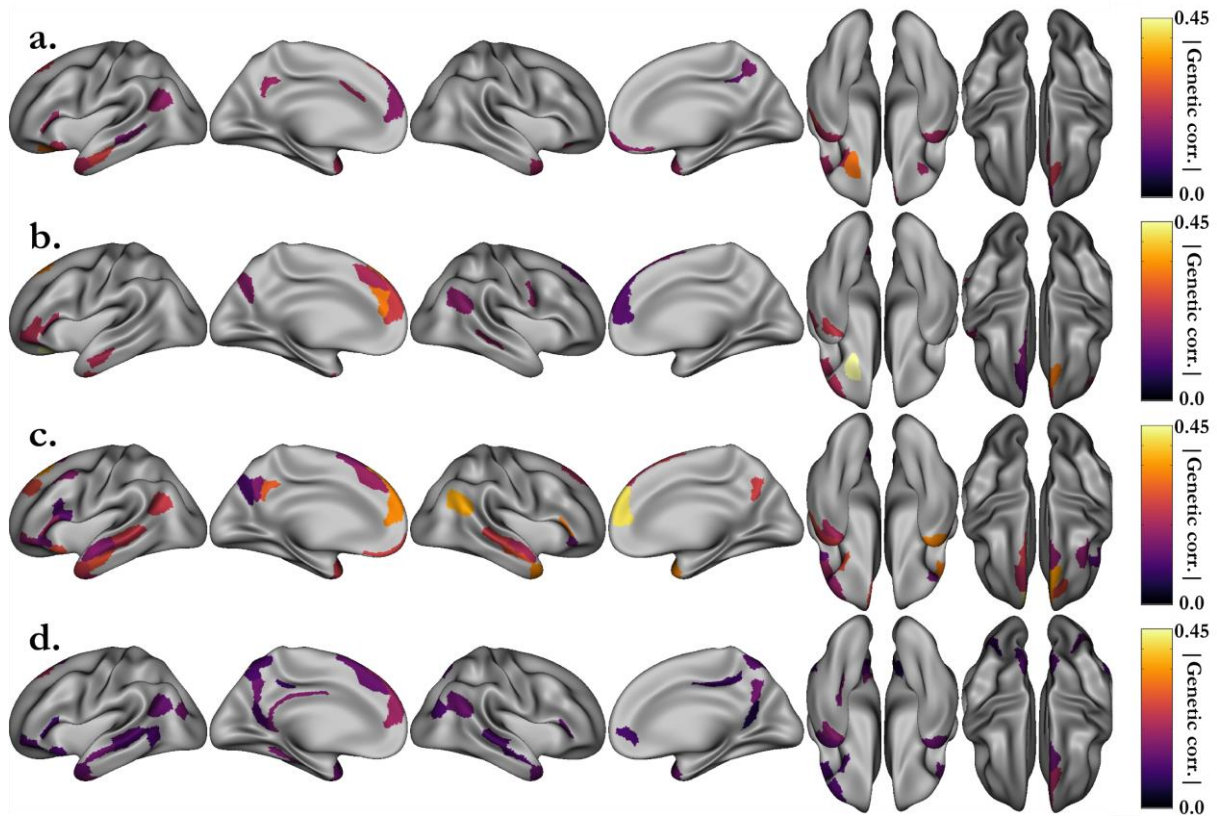


Figure 5. Shared genetic variance (absolute value) between the grayordinate activations of the STORY-MATH contrast and the NIH behavioral scores. a. Fluid intelligence (PMAT24_A_CR). **b.** Working memory (ListSort). **c.** Vocabulary comprehension (PicVocab). **c.** Reading decoding (ReadEng). Associated p-values ($p < 0.05$, uncorrected) can be found in Fig S4. Genetic correlation was investigated only for areals that were significantly phenotypically correlated (**Figure 4**).

Table 1. Heritability estimates for the behavioral scores associated with the tasks. Language accuracy and reaction time (RT) correspond to the average of the respective MATH and STORY variables. The p-values associated with the covariates related to age and sex, ethnic group and education level are also displayed.

Trait	h ² ±SE (p)	Age	Age ²	Sex	Age*Sex	Age ² *Sex	Hispanic	Educ	h ² cov%
		p-val							
Language Accuracy	0.34±0.06 (2.3·10 ⁻⁸)	0.49	0.75	0.01	0.96	0.22	0.97	1.2·10 ⁻⁷	7.2
Language RT	0.22±0.07 (8.4·10 ⁻⁴)	0.44	0.91	0.69	0.85	0.95	0.71	0.04	0.5
Math Accuracy	0.4±0.06 (1.6·10 ⁻¹⁰)	0.5	0.96	9.2·10 ⁻⁴	0.81	0.39	0.76	1.9·10 ⁻⁶	7.2
Math Difficulty Level	0.33±0.07 (1.0·10 ⁻⁶)	0.53	0.17	0.02	0.61	0.05	0.42	8.7·10 ⁻⁸	7.9
Math Median RT	0.17±0.07 (7.4·10 ⁻³)	0.6	0.41	0.78	0.82	0.69	0.26	0.01	1.1
Story Accuracy	0.18±0.06 (1.6·10 ⁻³)	0.78	0.48	0.86	0.76	0.23	0.61	2.2·10 ⁻³	1.4
Story Difficulty Level	0.33±0.07 (1.3·10 ⁻⁶)	0.66	0.43	0.81	0.44	0.28	0.91	0.11	0.0
Story Median RT	0.2±0.07 (1.4·10 ⁻³)	0.41	0.63	0.39	0.59	0.68	0.45	0.39	0.0