# Fluid and crystallised intelligence are associated with distinct regionalisation patterns of cortical morphology

C E Palmer[1], W Zhao[2], R Loughnan[2], C C Fan[1,3], W K Thompson[4], A M Dale[5,6,7] & T L Jernigan[1]

1. Center for Human Development, University of California, San Diego, 9500 Gilman Drive, La Jolla, CA 92161, USA
2. Department of Cognitive Science, University of California, San Diego, 9500 Gilman Drive, La Jolla, CA 92093, USA
3. Center for Multimodal Imaging and Genetics, University of California, San Diego School of Medicine, 9444 Medical Center Dr, La Jolla, CA 92037, USA
4. Division of Biostatistics, Department of Family Medicine and Public Health, University of California, San Diego, La Jolla, CA, USA
5. Department of Radiology, University of California, San Diego School of Medicine, 9500 Gilman Drive, La Jolla, CA 92037, USA
6. Department of Neuroscience, University of California, San Diego School of Medicine, 9500 Gilman Drive, La Jolla, CA 92037, USA
7. Department of Psychiatry, University of California, San Diego School of Medicine, 9500 Gilman Drive, La Jolla, CA 92037, USA

## ABSTRACT

Cognitive performance in children is predictive of academic and social outcomes; therefore, understanding neurobiological mechanisms underlying individual differences in cognition during development may be important for improving outcomes. Some theories of intelligence argue that a single latent, psychological construct with a specific neural substrate underlies many cognitive processes. In a previous study, we showed that a distributed configuration of cortical surface area and apparent thickness was associated with cognitive performance in a large sample (N=10,145) of nine and ten year old children from the Adolescent Brain and Cognitive Development (ABCD) study. Here we have compared the similarity of the configuration of cortical morphology best associated with cognitive performance across different tasks. We discovered strikingly distinct regionalisation patterns of cortical areal expansion and apparent thickness associated with measures of crystallised and fluid intelligence. The minimal overlap in these associations has important implications for competing theories about developing intellectual functions.

## INTRODUCTION

Intelligence is commonly defined using two broad components: fluid and crystallised[1,2]. Fluid intelligence refers to the ability to solve problems, reason, act quickly and adapt to novel situations[1,3]; whereas, crystallised intelligence encompasses task-specific knowledge that accrues throughout the lifespan. Although these factors are seemingly dissociable, they share common variance in hierarchical models of cognition. Indeed, the presence of a positive manifold in correlations of cognitive task performance is one of the most robust findings within psychology: individuals who perform well on a given cognitive task will tend to perform well on many other cognitive tasks[4,5]. The latent factor explaining this shared variance, termed 'g', is predictive of social and academic outcomes[6–8], and is highly heritable[7,9–11]. Previous studies have suggested that 'g' has unique structural and functional neural correlates[12–14]; however, these correlates vary greatly across studies, likely due to limited statistical power to identify replicable associations, differences in assessments used to estimate intelligence and differences in neuroimaging processing protocols[15]. However, the notion that the shared variance among test scores represents a causal *psychological* construct affecting general intelligence is heavily debated[16,17]. An alternative theory argues that the positive manifold develops through mutually beneficial interactions of independent cognitive processes during development[17,18]. In line with the latter theory, individual differences in cognitive performance would be associated with heterogeneous, distributed neural correlates.

Regional differences in cortical morphology during development (relative to global brain measures) have been previously linked to different aspects of cognitive function[19–23]. However, previous studies have been underpowered, and did not employ multivariate tools, to assess the contribution of the whole configuration of cortical morphology to individual differences in cognition. We have recently shown that the regionalisation of cortical surface area (CSA) and apparent cortical thickness (CTH), relative to global structural measures, was significantly associated with cognitive performance in a large sample (N=10,145) of 9 and 10 year old children from the Adolescent Brain Cognitive Development (ABCD) study[24]. In our previous study, we adopted the Multivariate Omnibus Statistical Test (MOSTest) to demonstrate that the aggregated effect across all cortical vertices was significantly associated with individual differences in cognitive performance. This was a particularly pertinent approach for understanding the relationship between regionalisation of cortical morphology and behaviour given the graded nature of the biology underlying the patterning of the cortex during embryonic development. These results support increasing evidence that many brain-behaviour associations are continuously and widely distributed across the cortical surface and highlight how individual differences in the patterning of cortical architecture may be important for explaining variability in cognition.

In the current study, we aimed to compare distributed patterns of associations between regional cortical morphology (CSA and CTH) and cognitive performance using the same sample from the ABCD study as Palmer and colleagues (2019). This large-scale study of 11,880 nine and ten year old children, uses neuroimaging, genetics and a multi-dimensional battery of behavioural assessments to investigate the role of various biological, environmental, and behavioural factors in brain, cognitive, and social/emotional development. Here we measured the similarity in the pattern of brain-behaviour associations across the cortex for the fluid and crystallised composite scores from the NIH Toolbox. Given the shared variance between these cognitve measures during development, we hypothesised that there would be a high degree of similarity between the cortical patterns of association for estimates of fluid and crystallised intelligence.

Importantly, the ABCD sample was recruited to resemble the population of the United States as closely as possible, therefore, the participants are from diverse racial, ethnic and socioeconomic backgrounds. There are large confounding associations between these demographic variables, brain structure and cognitive performance that complicate the interpretation of results. We have therefore repeated our analyses with and without controlling for these important sociodemographic variables in order to highlight the implications of this for the interpretation of our findings. This is an important aspect of analysing individual differences in large population-based samples such as the ABCD study.

# RESULTS

## Behavioural data

Figure 1A displays pairwise Pearson correlation coefficients describing the phenotypic relationship between all of the cognitive tasks and the composite scores measured at baseline in the ABCD study. As expected, performance across all tasks was positively correlated. Reading recognition and picture vocabulary performance were most highly correlated with each other (out of the single task measures; r=0.54). These scores were averaged to produce the crystallised composite score, therefore show a high correlation with this measure (r=0.85-0.89). The Toolbox measures used to produce the fluid composite score (flanker, dimensional card sorting, pattern processing speed, picture sequence and list working memory) were less correlated with each other (r=0.19-0.42), therefore showed slightly lower correlations with the fluid composite measure compared to the tasks contributing to the crystallised measure (r=0.62-0.7). The total composite score (mean of fluid and crystallised) was more highly correlated with the fluid measure (r=0.91) compared to the crystallised measure (r=0.79). The matrix reasoning task and little man task showed slightly higher correlations with the composite scores from the Toolbox compared to the Toolbox single task measures. This is unsurprising as the matrix reasoning task is often used as a measure of general intelligence and composite scores will show less measurement error. The RAVLT showed low correlations across all of the measures.

Figure 1B shows the same correlation matrix after pre-residualising all of the cognitive measures for the covariates of no interest (age, sex, race/ethnicity, household income and parental education) in order to show the phenotypic correlations after controlling for these confounding factors. Most of the correlation coefficients decreased in this analysis; however, the overall pattern of these relationships remained consistent.

## Distinct patterns of association between the regionalisation of CSA and the fluid and crystallised composite scores

In a previous study, we demonstrated significant associations between the fluid and crystallised composite scores (measured in the same sample) and the regionalisation of CSA and CTH using a novel multivariate omnibus test: the MOSTest[24]. The MOSTest aggregates effects across the entire cortex and therefore is better able to detect associations that are distributed across the cortex compared to a standard univariate neuroimaging omnibus test that assumes effects are sparse and localised. These results showed a significant association between the regionalisation of CSA and CTH and these composite scores as well as each of the single task measures averaged to produce these composite scores.

In the current study, we aimed to measure the similarity across the estimated effect size maps for these associations. Interestingly, on visualising the estimated effect size maps of the associations between the regionalisation of CSA (controlling for total CSA) and the fluid and crystallised composite scores, we see a unique structural pattern of association for each composite score (Fig 2A&B). Figure 2C shows the difference between the beta coefficients for the fluid (F) and crystallised (C) surface maps, which demonstrates the difference in the distributed pattern of effects. To quantify the magnitude of these vertex-wise differences relative to the original associations, we calculated a ratio of the variance (root mean squared) of the beta coefficient differences divided by the variance (root mean squared) of the average beta coefficients across the F and C estimated effect size maps (RMS ratio for F – C = 1.21).

In order to determine the unique pattern of association between the relative configuration of CSA and the fluid and crystallised composite scores, we generated a vertex-wise association map between the regionalisation of CSA and the fluid composite score controlling for the crystallised composite score ($F_C$; figure 2E), and the crystallised composite score controlling for the fluid composite score ($C_F$; figure 2H). The pattern of associations was almost identical to that obtained when the respective estimated intelligence measure was not included as a covariate in the model. Indeed, the surface map of the vertex-wise differences between the associations for each composite score and that composite score controlling for the other were very small in magnitude compared to the original F - C effect size maps (figure 2F,I; RMS ratio for C – $C_F$ = 0.27; RMS ratio for F – $F_C$ = 0.38).

Moreover, the vertex-wise correlation between the estimated beta coefficients for F vs $F_C$ was high (r=0.92) as was the correlation between the beta coefficients for C vs $C_F$ (r=0.95). In contrast, the vertex-wise correlation between the beta coefficients for F vs C was much lower (r=0.30). This further implies that there was minimal shared variance between the associations for the fluid and crystallised scores and the regionalisation of CSA. The minimal overlap in the cortical configuration associated with these measures supports that the configurations of relative CSA associated with these measures were relatively distinct.

## Distinct patterns of association between the regionalisation of CTH and the fluid and crystallised composite scores

Estimated effect size maps showing the vertex-wise associations between the relative configuration of CTH and the fluid and crystallised composite scores are shown in figure 3A&B. The structural pattern of association between these composite scores and CTH were more similar than with CSA. However, there were still key regions with distinct differences in relative CTH associations between the fluid and crystallised composite scores. Figure 3C shows the difference between the beta coefficients for the F and C surface maps, which were relatively large compared to the original estimated effect sizes (RMS ratio for F – C = 0.92). In order to determine the unique pattern of association between the regionalisation of CTH and F and C, we generated a vertex-wise map of the association between the regionalisation of CTH and $F_C$ (figure 3E) and CTH and $C_F$ (figure 3H). As with CSA, the pattern of associations was almost identical to when the respective intelligence measure was not included in the model. Indeed, the estimated effect size map of the vertex-wise differences between the associations for each composite score and each composite score controlling for the other were very small in magnitude (figure 3F,I; RMS ratio for C – $C_F$ = 0.31; RMS ratio for F – $F_C$ = 0.30). The vertex-wise correlations between the map of associations for each composite score and that composite score controlling for the other were high (F vs $F_C$: r=0.94; C vs $C_F$: r=0.95); whereas the vertex-wise correlation between the maps of association for the fluid and crystallised composite scores was much lower (F vs C: r=0.40). As with CSA, the minimal overlap in the cortical configurations

associated with these measures supports that the configurations of CTH associated with these measures were relatively distinct.

We previously showed that the association between relative CTH and cognition was related to total CSA; therefore, we repeated these analyses including total CSA as an additional covariate in the mass univariate GLM. The estimated effect size maps for the crystallised and fluid composite scores were more correlated, but remained relatively distinct as shown by the magnitude of the beta coefficient differences (RMS ratio for F – C = 0.95; C – $C_F$ = 0.38; F – $F_C$ = 0.24) and the vertex-wise effect size correlations (F vs C: r=0.55; F vs $F_C$: r=0.97; C vs $C_F$: r=0.93). The estimated effect size maps for the association between relative CTH (controlling for total CSA) and each of the cognitive tasks can be found in supplementary figure 1.

## Distinct patterns of association for the regionalisation of CTH and CSA across cognitive tasks

Visualising the effect size maps between each cortical measure and each cognitive task revealed relatively distinct patterns of association across the cognitive tasks (figure 4). This was further supported by the low vertex-wise correlations across these associations (figure 5). The estimated beta coefficients for the regionalisation of CTH and cognition were more correlated across tasks than for the regionalisation of CSA and cognition. The single tasks showed lower correlations between them and high correlations with the relevant composite measures. Interestingly, the effect size maps appeared to be more similar for the tasks using similar underlying cognitive processes, such as for the picture vocabulary task and oral reading recognition task. Moreover, the composite measures appeared to reflect mixtures of the patterns of association for each of the tasks that were averaged to produce the composite scores.

We then generated the estimated effect size maps for each cortical measure and each cognitive task and composite score without controlling for the sociodemographic variables of race/ethnicity, household income and parental education. The vertex-wise correlations between the estimated beta coefficients for each of the cognitive tasks predicted by the regionalisation of CSA are shown in figure 5C and by the regionalisation of CTH are shown in figure 5D. All of the correlation coefficients across tasks increased compared to the correlation matrices when controlling for the sociodemographic factors. This supports that some of the variance in the sociodemographic factors is shared with the cognitive and structural measures.

## Association maps when not controlling for sociodemographic variables

We computed the estimated effect size maps for the F and C composites scores associated independently with the regionalisation of CSA (figure 6A&B) and CTH (figure 6D&E) without controlling for the sociodemographic variables. For both morphology measures, the vertex-wise correlation between the F and C maps increased (CSA: r=0.69; CTH: r=0.89). Interestingly, however, the F-C difference maps were remarkably similar to when the sociodemographic variables were controlled for as shown by a vertex-wise correlation between the F-C difference maps with and without controlling for sociodemographic factors (CSA: r=0.96; CTH: r=0.94). This strongly suggests that these confounding variables in aggregate index common shared variance between cortical morphology and cognition across domains that is not specific to a particular task performance. Estimated effect size maps for the association between the regionalisation of CSA and CTH and cognitive performance across tasks without controlling for sociodemographic factors can be found in supplementary figure 2.

Partitioning the variance between brain structure and cognition

In our previous study, we used a Bayesian polyvertex score ($PVS_B$) to quantify the variance in each cognitive phenotype predicted by the vertex-wise imaging data for the regionalisation of CSA and CTH (see supplementary figure 3 for a summary of these findings[24]). The $PVS_B$ represents a linear weighted sum of the vertex-wise associations, which are projected from a training set to an independent hold-out set within a cross validation framework. Importantly, the values reported previously represented a conservative, lower bound of the variance explained when potentially confounding variables were controlled for. Here, we have calculated the behavioural variance explained when these sociodemographic variables were not controlled for. We also calculated the variance in cognitive performance predicted by the sociodemographic variables without accounting for brain structure to understand the proportion of individual variability currently accounted for by these measures. All of these values represent an out of sample $R^2$ calculated using a 10-fold cross validation framework.

A series of models were estimated to quantify and partition the shared and unique variance in cognitive performance predicted by cortical morphology and the sociodemographic factors. The results are summarised in Figure 7. Without controlling for brain structure, the sociodemographic factors collectively predicted more variance in crystallised (25.17%$R^2$) compared to fluid scores (13.46%$R^2$). Across regionalisation measures, the structural brain phenotypes explained ~30% of the association between the sociodemographic variables and the crystallised scores and ~20% of the association between the sociodemographic variables and the fluid scores.

The regionalisation of CSA explained a smaller proportion of the total variability in cognition compared to the regionalisation of CTH (for fluid scores: total CSA = 2.06%$R^2$; CSA $PVS_B$ = 2.39%$R^2$; CTH $PVS_B$ =4.82%$R^2$; for crystallised scores: total CSA predicted 5.13%$R^2$; CSA $PVS_B$ predicted 3.33%$R^2$; CTH $PVS_B$ = 8.02%$R^2$). Mean CTH was not predictive of cognitive performance. Additionally controlling for the sociodemographic factors lead to a decrease in the magnitude of individual variability in the composite scores explained by brain structure. Across imaging measures the sociodemographic variables accounted for ~70-90% of the variation in fluid and crystallised scores predicted by brain structure. This shows that the relationship between brain structure and cognition was strongly related to the sociodemographic factors.

Importantly, we previously showed that some of the variability in cognition explained by relative CTH was associated with total CSA. Therefore, we also calculated the variance in these composite scores explained by the regionalisation of CTH after additionally controlling for total CSA. Controlling for total CSA reduced the variance in cognition predicted by the regionalisation of CTH. The individual variability in fluid scores predicted by the regionalisation of CTH controlling for total CSA was 3.00%$R^2$ and for the crystallised scores was 3.38%$R^2$. Controlling for sociodemographic factors reduced these estimates by 70-86%.

# DISCUSSION

In this study, we have discovered distinct patterns of association between the regionalisation of cortical morphology and cognitive test performance in a large sample of 9 and 10 year old children (N=10,145). In particular, individual differences in estimates of fluid and crystallised intelligence were associated with distinct maps of relative cortical areal expansion and apparent thickness. This suggests that regional cortical architecture is differentially related to these two types of cognitive function. The relatively distinct association maps for the single tasks used to

generate these composite measures provides evidence in support of mutualism or sampling models of intelligence, which argue that no single, latent neural construct is required to account for the strong positive manifold underlying "g". Furthermore, we have shown that the sociodemographic diversity within ABCD impacts the association between cortical morphology and cognition similarly across cognitive domains. Since the treatment of sociodemographic variables in these models led to substantial differences in the magnitude of effects of interest, users of ABCD data should select covariates carefully and discuss any sociodemographic factors potentially related to both explanatory variables and outcomes.

## Distinct associations between cortical morphology and different cognitive modalities

We recently demonstrated that there were significant associations between the regionalisation of CSA and CTH (when controlling for global structural measures) and the fluid and crystallised composite scores from the NIH Toolbox within the ABCD baseline sample using the Multivariate Omnibus Statistical Test: the MOSTest [24]. This test aggregates vertex-wise associations across the cortex thereby improving our power for detecting widely distributed effects compared to the standard neuroimaging approach[24–26]. Given the graded nature of the biology underlying the patterning of the cortex this is a more appropriate statistical approach compared to standard neuroimaging methods. As a secondary analysis, in the current study, we sought to determine the similarity in the cortical configurations associated with these composite measures. A surprising result was the striking difference in the patterns observed for the fluid and crystallised scores. This has important implications for understanding the development of cognition in childhood.

The estimated effect size maps between the regionalisation of CSA and the crystallised composite score (comprised of reading and vocabulary measures) showed that greater areal expansion in regions previously associated with language functions, such as the left temporal and middle frontal regions[27–29], was associated with higher crystallised scores; whereas, greater areal expansion in a different set of regions, previously implicated in cognitive control mechanisms required for the fluid tasks, such as the anterior cingulate and insula cortices[19,20,30–32], was associated with higher fluid scores. The patterns of apparent CTH associated with these measures were slightly less distinct, however to some degree appeared to mirror these associations with areal expansion. For both morphology measures, there was a clear pattern of positive and negative associations with cognitive performance, which suggests that, at this developmental stage, *relative* differences in the size and thickness of cortical regions mirror individual difference variability in cognition.

Interestingly, the areas most strongly associated with the total composite score are a combination of those most strongly associated with either fluid or crystallised intelligence. The maps of association between both the regionalisation of CSA and CTH and the total composite score represent a mixture of the fluid and crystallised associations. Indeed, the same can be seen for the fluid and crystallised composite score maps, which appear to represent a mixture of the association maps for each of the individual tasks that contributed to those scores. This is reflected within the correlation matrix of the vertex-wise associations (figure 5A&B). This observation has important implications for theories of intelligence.

## Theories of intelligence: sampling, mutualism or 'g'?

Factor analyses of cognitive measures frequently reveal a higher order general latent factor 'g' that can explain (statistically) individual differences in cognitive performance. However, this observation alone does not necessarily reveal anything about the origins or development of 'g'.

Here we clearly do not observe a single, causal neuroanatomical substrate that may represent 'g' (i.e., influence performance on all cognitive tasks). Brain size is frequently associated with "general cognitive ability" and thought to represent 'g'. Indeed total CSA was modestly associated with cognitive performance in the current study; however, before controlling for demographic factors, total CSA only predicted ~5% of the variability in crystallised scores and ~2% in fluid scores. Due to the unprecendented statistical power in this study, we can infer that total CSA is unlikely, across other studies, to account for any more individual variability in cognitive scores in children of this age. Indeed, this effect is much smaller than the shared variance across cognitive measures that has been attributed to 'g' (~21%) in a similar sample from the ABCD study[33]. Moreover, the relative importance of this measure for predicting individual variability in cognition differed as a function of the type of intelligence measured, which is inconsistent with this being a global effect. These results do not rule out that a single or global neural phenotype underlying intelligence could be represented within a different neural modality, such as functional connectivity or latency of neuronal signalling; however, there is a lack of supporting evidence for a dominant neural 'g' phenotype in the current study. There are alternative models of intelligence that attempt to explain how the positive manifold (the positive correlation amongst cognitive tasks) can be generated from multiple, independent psychological constructs, which fit more in line with the results shown here.

The sampling model posits that because cognitive tasks do not measure a single cognitive process but instead require multiple processes, the overlap in the processes required across similar tasks produces a positive correlation between them[34,35]. An alternative mutualism model simulates the positive manifold via mutually beneficial interactions between initially uncorrelated systems that support cognitive behaviours[17,18]. For example, having good short-term memory may aid the development of cognitive strategies, which can in turn lead to better short-term memory [36]. In this model, each initially independent system has a biological constraint, or resource, which limits the potential growth of that system. These resources may be determined by either genetics or the environment, or both, and confer individual differences in how these systems develop and interact over time. The weighted sum of these resources, in simulations using the mutualism model, is a good predictor of cognitive performance and correlates highly with the 'g' that emerges from mutualism. This is an important distinction from the more conventional 'g' model as here individual differences in mature cognitive performance emerge from different combinations of these resources. For example, the same IQ score could arise from high memory resources in one individual, but high processing speed in another individual.

In the current study, the configuration of cortical morphology associated with the total composite score (our proxy for 'g') appeared to be a mixture (weighted sum) of structural associations with the other cognitive tasks. In turn, each of these tasks clearly does not represent a single cognitive process, but requires a different combination of many cognitive processes. Here we suggest that relative differences in cortical morphology may reflect to some degree the scaffolding, or resources, contributing to the performance of these processing systems, and different mixtures of these configurations may map to different levels of performance on each cognitive task.

## Understanding individual variability in cognition related to both brain structure and sociodemographic factors

Before controlling for demographic variables, the proportion of variance in the composite measures of cognitive performance shared with the brain phenotypes was ~8% for the crystallised scores and ~5% for the fluid scores. Overall, these phenotypes and the partially

confounded sociodemographic variables together accounted statistically for ~34% of the variance in crystallised scores and ~21% of the variance in fluid scores. These results suggest that many factors that relate to individual differences in cognitive test scores remain unaccounted for in these statistical models. However, the results are sufficiently powerful to constrain future hypotheses by restricting the range of plausible causal effects of well-measured variables, such as total CSA.

The cortical arealisation phenotype and regionalisation of apparent CTH both accounted for additional unique variance in cognition independent of total CSA: ~3% for crystallised and ~2% for fluid scores. Even after pre-residualising for sociodemographic factors that could index cultural, environmental and other experiential effects on cognitive test scores, as well as global brain measures, the associations between the cortical regionalisation phenotypes and cognitive test scores were statistically robust; but only explained ~0.5% of the residual variation. These effects may be very small, but that is perhaps not surprising given the number of factors that can influence cognition. Indeed, our results show that the sociodemographic variables accounted for a substantial proportion of the shared variance between the brain phenotypes and composite measures of cognitive function. This may result from many small effects of factors indexed by these variables, such as nutrition, limited access to healthcare and education, untreated prenatal complications, high levels of stress or exposure to environmental toxins.

It is possible that the dissociable, specific variance across cognitive domains associated with cortical regionalisation, after controlling for sociodemographic factors, is mediated by genetic variability in cortical arealisation. Indeed, the regionalisation of the cortex is moderately heritable[37–39] and during early embryonic development, gradients of morphogens dictate the expression of transcription factors across the dorsal proliferative zone, which determines the eventual spatial location and functional specialisation of cortical projection neurons[40,41]. Such biases in arealisation may influence developing cognitive functions in an individual in ways that advantage some functional domains relative to others, which creates important diversity within our population. The brain-behaviour associations in the current study most likely reflect both genetically mediated differences in cortical architecture and both independent and correlated interactions with the environment and other developing brain systems. Indeed, within our education system, particular individuals may be selectively advantaged and thus may be disproportionately exposed to environments that can enhance the positive interactions between developing systems, multiplying a specific genetic effect. This gene-environment correlation can inflate heritability estimates and increase individual variability in estimates of general intelligence [42–45].

Given our current findings, we hypothesise that it is unlikely that there is a single cortical configuration underlying general intelligence; however, there may be a dominant phenotype, or configuration of resources (shown here as the mean association pattern), that may offer an advantage in our society and be selected for by our education system. Subsequent analyses are required to directly test this hypothesis by quantifying the heterogeneity in cortical architecture across this sample and determining if there are different clusters of individuals within the sample with different cortical configurations that are associated with the same cognitive performance. This will provide increased support for mutualism models of intelligence in which variability in the configuration (weighting) of resources across individuals can result in similar levels of general intelligence. This may also explain why the mean association maps explain a very small proportion of variability in behaviour.

## Conclusions

The graded and distributed nature of the neurobiology underlying the regionalisation of the cortex supports the importance of studying the whole configuration of cortical morphology associated with behaviour using multivariate statistics and continuous pattern comparison methods. Using the unprecedented ABCD dataset, we now have the power to discover novel brain-behaviour associations and understand individual variability in behaviour among a diverse sample of participants. With future releases of the ABCD data we will track how the regionalisation of CSA and CTH and cognitive performance change over time within an individual to further our understanding of these relationships.

## Online Methods

### *Sample*

The ABCD study is a longitudinal study across 21 data acquisition sites following 11,878 children starting at 9 and 10 years old. This paper analysed the full baseline sample from release 2.0.1 (NDAR DOI: 10.15154/1504041). The ABCD Study used school-based recruitment strategies to create a population-based, demographically diverse sample, however it is not necessarily representative of the U.S. national population [46,47]. Due to the inclusion of a wide range of individuals across different races, ethnicities and socioeconomic backgrounds, it is important to carefully consider how to control for these potentially confounding factors and the implications of this on our effects of interest. Sex and age were used as covariates in all analyses.

Only subjects who had complete data across all of the measures analysed were included in the neuroimaging analyses. There was a large number of subjects with missing income data (n=1018), therefore missing values were imputed by taking the median income value across participants from the same testing site. The sources of missing data were as follows: incomplete across all demographic variables (n=189), incomplete across all cognitive measures (n=944), unavailable T1-weighted MRI scan for reasons outlined in the ABCD release notes (e.g. did not get scanned, motion artefacts) (n=339) and imaging data that was made available but did not pass the free-surfer QC flag (n=462). These missing data values are not mutually exclusive. The permutation testing used Palmer et al (2019) was dependent on having multiple families with the same number of children, therefore the single family with 5 children was excluded from these analyses in order to match the final sample in our original study. This resulted in a final sample of 10,145 subjects. Supplementary Table 1 displays the names of each variable used in these analyses from data release 2.0.1. Supplementary Table 2 shows the demographic characteristics of the sample as a function of cognitive performance on the NIH Toolbox total composite score.

## Neurocognitive assessment

### *NIH Toolbox Cognition Battery®*

The NIH Toolbox Cognition Battery® (NTCB) is a widely used battery of cognitive tests that measures a range of different cognitive domains. All of the tasks within the NTCB were administered using an iPad with support or scoring from a research assistant where needed. Below is a brief description of each task.

The ***Toolbox Oral Reading Recognition Task® (TORRT)*** measured language decoding and reading. Children were asked to read aloud single letters or words presented in the center of an iPad screen. The research assistant marked pronunciations as correct or incorrect. Extensive training was given prior to administering the test battery. Item difficulty was modulated using computerised adaptive testing (CAT).

The ***Toolbox Picture Vocabulary Task® (TPVT)***, a variant of the Peabody Picture Vocabulary Test (PPTV), measured language and vocabulary comprehension. Four pictures were presented on an iPad screen as a word was played through the iPad speaker. The child was instructed to point to the picture, which represented the concept, idea or object name heard. CAT was implemented to control for item difficulty and avoid floor or ceiling effects.

The ***Toolbox Pattern Comparison Processing Speed Test® (TPCPST)*** measured processing speed. Children were shown two images and asked to determine if they were identical or different by touching the appropriate response button on the screen. This test score is the sum of the number of items completed correctly in the time given.

The ***Toolbox List Sorting Working Memory Test® (TLSWMT)*** measured working memory. Children heard a list of words alongside pictures of each word and were instructed to repeat the list back in order of their actual size from smallest to largest. The list started with only 2 items and a single category (e.g. animals). The number of items increased with each correct answer to a maximum of seven. The child then progressed to the next stage in which two different categories were interleaved. At this stage children were required to report the items back in size order from the first category followed by the second category. Children were always given two opportunities to repeat the list correctly before the experimenter scored the trial as incorrect.

The ***Toolbox Picture Sequence Memory Test® (TPSMT)*** measured episodic memory. On each trial, children were shown a series of fifteen pictures in a particular sequence. The pictures illustrated activities or events within a particular setting (e.g. going to the park). Participants were instructed to touch the pictures in the original sequence in which they were shown. The Rey-Auditory Verbal Learning Task (RAVLT) was also included in the ABCD neurocognition battery as a more comprehensive measure of episodic memory.

The ***Toolbox Flanker Task® (TFT)*** measured executive function, attentional and inhibitory control. This adaptation of the Eriksen Flanker task [48] captures how readily a participant is influenced by the congruency of stimuli surrounding a target. On each trial a target arrow was presented in the center of the iPad screen facing to the left or right and was flanked by two additional arrows on both sides. The surrounding arrows were either facing in the same (congruent) or different (incongruent) direction to the central target arrow. The participant was instructed to push a response button to indicate the direction of the central target arrow. Accuracy and reaction time scores were combined to produce a total score of executive attention, such that higher scores indicate a greater ability to attend to relevant information and inhibit incorrect responses.

The ***Toolbox Dimensional Change Card Sort Task® (TDCCS)*** measured executive function and cognitive flexibility. On each trial, the participant was presented with two objects at the bottom of the iPad screen and a third object in the middle. The participant was asked to sort the third object by matching it to one of the bottom two objects based on either colour or shape. In the first block participants matched based on one dimension and in the second block they switched to the other dimension. In the final block, the sorting dimension alternated between trials pseudorandomly. The total score was calculated based on speed and accuracy.

In the current study, the uncorrected scores for each task were used for statistical analyses. Composite scores of crystallised intelligence (mean of TPVT and TORRT), fluid intelligence (mean of TPCPST, TLSWMT, TPSMT, TFT and TDCCS) and total score (mean of all tasks) were also analysed. These measures are highly correlated with 'gold standard' measures of intelligence in adults [49] and children [50].

*Rey-Auditory Verbal Learning Task (RAVLT)*

This task measures auditory learning, recall and recognition. Participants listened to a list of 15 unrelated words and were asked to immediately recall these after each of five learning trials. A second unrelated list was then presented and participants were asked to recall as many words

as possible from the second list and then recall words again from the initial list. Following a delay of 30 minutes (during which other non-verbal tasks from the cognitive battery are administered), longer-term retention was measured using recall and recognition. This task was administered via an iPad using the Q-interactive platform of Pearson assessments [51]. In the current study, the total number of items correctly recalled across the five learning trials was summed to produce a measure of auditory verbal learning.

*Little Man Task (LMT)*

This task measures visuospatial processing involving mental rotation with varying degrees of difficulty [52]. A rudimentary male figure holding a briefcase in one hand was presented on an iPad screen. The figure could appear in one of four positions: right side up vs upside down and either facing the participant or with his back to the participant. The briefcase could be in either hand. Participants indicated which hand the briefcase was in using one of two buttons. Performance across the 32 trials was measured by the percentage of trials in which the child responded correctly. This was divided by the average reaction time to complete the task (in seconds) to produce a measure of efficiency of visuospatial processing. This was the dependent variable analysed in this study.

*Matrix reasoning*

Nonverbal reasoning was measured using an automated version of the Matrix Reasoning subtest from the Weschler Intelligence Test for Children-V (WISC-V; Weschler, 2014). On each trial the participant was presented with a series of visuospatial stimuli, which was incomplete. The participant was instructed to select the next stimulus in the sequence from four alternatives. There were 32 possible trials and testing ended when the participant failed three consecutive trials. The total raw score, used in the current study, was the total number of trials completed correctly.

## MRI acquisition

The ABCD MRI data were collected across 21 research sites using Siemens Prisma, GE 750 and Philips 3T scanners. Scanning protocols were harmonised across sites. The T1w acquisition (1 mm isotropic) was a 3D T1w inversion prepared RF-spoiled gradient echo scan using prospective motion correction, when available [54,55] (echo time = 2.88 ms, repetition time = 2500 ms, inversion time = 1060 ms, flip angle = 8°, FOV = 256x256, FOV phase = 100%, slices = 176). Only the T1w scans were analysed in this paper. Full details of all the imaging acquisition protocols used in ABCD are outlined by Casey et al [56]. Scanner ID was included in all analyses to control for any differences in image acquisition across sites and scanners.

## Image pre-processing

Pre-processing of all MRI data for ABCD was conducted using in-house software at the Center for Multimodal Imaging and Genetics (CMIG) at University of California San Diego (UCSD) as outlined in Hagler et al [57]. Manual quality control was performed prior to the full image pre-processing and structural scans with poor image quality as well as those that did not pass FreeSurfer QC were excluded from all analyses. Brain segmentation and cortical surface reconstruction were completed using FreeSurfer v5.3.0 [58,59].

T1-weighted structural images were corrected for distortions caused by gradient nonlinearities, coregistered, averaged, and rigidly resampled into alignment with an atlas brain. See previous publications for details of the surface based cortical reconstruction segmentation procedures [58–62]. In brief, a 3D model of the cortical surface was constructed for each subject. This included segmentation of the white matter (WM), tessellation of the gray matter (GM)/WM boundary, inflation of the folded, tessellated surface, and correction of topological defects.

Measures of cortical thickness at each vertex were calculated as the shortest distance between the reconstructed GM/WM and pial surfaces [60]. To calculate cortical surface area, a standardised tessellation was mapped to the native space of each subject using a spherical atlas registration, which matched the cortical folding patterns across subjects. Surface area of each point in atlas space was calculated as the area of each triangle. This generated a continuous vertex-wise measure of relative areal expansion or contraction. Cortical maps were smoothed using a Gaussian kernel of 20 mm full-width half maximum (FWHM) and mapped into standardised spherical atlas space. Vertex-wise data for all subjects for each morphometric measurement were concatenated into matrices in MATLAB R2017a and entered into general linear models for statistical analysis using custom written code.

## Statistical Analysis

*Vertex-wise effect size maps*

All behavioural variables were standardised (z-scored) prior to analysis as was the vertex-wise brain data. We applied a general linear model (GLM) associating a given behaviour $y$ from a set of covariates $W$ and the vertex-wise morphology data, $X_v$

$$y = \alpha_{0,v} + \boldsymbol{W}\alpha_v + X_v\beta_v + \varepsilon_v$$

Here, $W$ represents a standardized $N\ x\ m$ matrix of covariates of no interest where $N$ represents the number of subjects and $m$ the number of covariates. $X_v$ denotes the standardised $N\ x\ 1$ vector of imaging data for the $v$th vertex. This GLM was applied univariately at each vertex $v = 1, \ldots V$ . Let $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_V)'$ and α denote the $V\ x\ 1$ and $m\ x\ 1$ vectors of parameters of interest and no interest, respectively. For visualisation of these effects the standardised $\boldsymbol{\beta}$ coefficients were plotted across vertices.

This analysis was repeated for all of the thirteen cognitive measures predicted by relative CSA (controlling for total CSA) and relative CTH (controlling for mean CTH) separately resulting in N=26 independent vertex-wise analyses. Models determining the association between CSA and behaviour included total CSA (sum of CSA across vertices) as an additional predictor in the design matrix, $W$, and models analysing CTH included the mean CTH across vertices as an additional predictor. Mass univariate standardised beta surface maps were created showing the vertex-wise associations for each analysis. These analyses were also repeated without including race/ethnicity, income and parental education as predictors in the covariates matrix, $W$.

We previously found that the regionalisation of CTH was not completely independent of total CSA; therefore, we repeated the analyses for relative CTH including total CSA as an additional covariate.

*Comparison across associations maps*

To determine the shared variance between the effect size maps for the fluid (F) and crystallised (C) composite scores for each imaging phenotype, additional maps were created controlling for the other composite measure by including that measure as an additional covariate within the design matrix, $W$. This produced a vertexwise standardised beta effect size surface map for F independent of the association between cortical morphology and C ($F_C$) and C independent of F ($C_F$). Surface maps of the difference between the vertexwise beta coefficients as well as vertexwise Pearson correlation coefficients were calculated for the following contrasts: F – C, F – $F_C$ and C – $C_F$. The magnitude of effects across the cortex for each map was calculated using the following formula: $= \sqrt{\overline{\beta^2}}$. In order to determine the relative size of the effects across the difference maps compared to the individual maps we calculated an RMS ratio using the following formula: $RMS_{DIFF}/(\frac{RMS_1+RMS_2}{2})$, with $RMS_{DIFF}$ representing the RMS of the difference map and RMS1 and RMS2 as the RMS for each individual map used to produce the difference.

*Quantifying the magnitude of the association between brain structure and cognition using a Bayesian PVS ($PVS_B$)*

The $PVS_B$ was calculated for the fluid and crystallised composite scores to quantify the behavioural variance explained by the vertex-wise cortical morphology. All behavioural and imaging data were pre-residualised using the covariates of no interest prior to calculation of the $PVS_B$. We completed this step multiple times including different covariates in order to determine the change in $R^2$ when particular demographic variables were and were not controlled for.

For the imaging data, we included the global CSA and CTH measures specific to each modality in the pre-residualisation as covariates. This allowed us to determine the unique association between relative cortical morphology and cognition and compare this to the predictive power of brain structure without controlling for global measures. The method used here is outlined in detail by Zhao and colleagues [25]. The association between each imaging phenotype and each cognitive task was modelled using the mass univariate approach such that the behaviour of interest was predicted independently at each vertex using a general linear model (GLM). A Bayesian posterior mean effect size was then calculated at each vertex to take into account the correlation amongst vertices. This method was adapted from the LDpred framework originally developed to improve accuracy for polygenic risk scores [63]. The PVS predicting behaviour from cortical morphometry was then computed for each subject by as the product sum of the estimated Bayesian effects and the pre-residualized cortical morphometry vector. This measure thus harnesses the explanatory power of all of the vertices with respect to behaviour. The computed PVS is then be compared with the observed behaviour in order to provide an estimate for how much variance in the observed behaviour can be predicted using the vertex-wise imaging phenotype.

In order to generate an unbiased PVS-B for every subject we used a leave-one-out 10 fold cross-validation procedure. The Bayesian parameters were estimated in the training set (90% of full sample) and multiplied with the imaging phenotype of participants in the test set (10% of full sample). This was repeated 10 times for each fold until a PVS-B was calculated for every participant in the full sample. The subjects in each fold were randomly selected based on unique family IDs, such that subjects within the same family were always within the same fold. The association between the imaging phenotype and behaviour across the whole sample was calculated as the squared correlation ($R^2$) between the observed behaviour and the predicted

behaviour (the PVS-B). This process was repeated for four imaging phenotypes: CSA, relative CSA (controlling for total CSA), CTH. relative CTH (controlling for mean CTH).

In order to determine the unique variance in cognitive performance predicted by relative CTH independent of total CSA, we calculated another $PVS_B$ for relative CTH including total CSA as a covariate in the pre-residualisation of the imaging data for each of the cognitive tasks.

In order to explore the proportion of shared variability in cognitive performance explained by brain structure and the demographic variables, we generated several linear models with differing predictors. These models were generated separately for each imaging modality (when these were included) and with either the fluid or crystallised scores as the dependent variable. Each model was trained on 90% of the sample and tested in a 10% hold out set within a 10-fold cross validation framework to produce a robust, out-of-sample $R^2$. The models are outlined in figure 7. Where indicated both the dependent and independent variables were pre-residualised for the stated variables (in parentheses), in order to remove the variance associated with those variables and calculate the unique variance associated only with the variables of interest. This method was used to partition the shared and unique variance across the measures of interest. It is important to note, this method only offers an approximation of the unique behavioural variance associated with these variables of interest, because these variables are not completely orthogonal to one another, therefore this variance is dependent upon the order in which these variables are pre-residualised. Here we always pre-residualise for age, sex and scanner first as nuisance variables except for in the full model. The full model (model 1) gives an estimate of the maximal variance explained when all of the variables measured are included in a model together and thus their covariance is accounted for when estimating the model $R^2$.

## Acknowledgements

The authors wish to thank the youth and families participating in the Adolescent Brain Cognitive Development (ABCD) Study and all ABCD staff. Data used in the preparation of this article were obtained from the Adolescent Brain Cognitive Development (ABCD) Study (https://abcdstudy.org), held in the NIMH Data Archive (NDA). This is a multisite, longitudinal study designed to recruit more than 10,000 children age 9-10 and follow them over 10 years into early adulthood. The ABCD Study is supported by the National Institutes of Health and additional federal partners under award numbers U01DA041022, U01DA041028, U01DA041048, U01DA041089, U01DA041106, U01DA041117, U01DA041120, U01DA041134, U01DA041148, U01DA041156, U01DA041174, U24DA041123, U24DA041147, U01DA041093, and U01DA041025. A full list of supporters is available at https://abcdstudy.org/federal-partners.html. A listing of participating sites and a complete listing of the study investigators can be found at https://abcdstudy.org/Consortium_Members.pdf. ABCD consortium investigators designed and implemented the study and/or provided data but did not all necessarily participate in analysis or writing of this report. This manuscript reflects the views of the authors and may not reflect the opinions or views of the NIH or ABCD consortium investigators. The ABCD data repository grows and changes over time. The data was downloaded from the NIMH Data Archive ABCD Collection Release 2.0.1 (DOI: 10.15154/1504041).

## REFERENCES


1. Horn, J. L. & Cattell, R. B. Refinement and test of the theory of fluid and crystallized general intelligences. *J. Educ. Psychol.* **57**, 253–270 (1966).

2. Deary, I. J. Intelligence. *Annu. Rev. Psychol.* **63**, 453–482 (2012).

3. Salthouse, T. A. When does age-related cognitive decline begin? *Neurobiol. Aging* **30**, 507–514 (2009).

4. Spearman, C. "General Intelligence," Objectively Determined and Measured. *Am. J. Psychol.* **15**, 201 (1904).

5. Wechsler, D. *The measurement of adult intelligence (3rd ed.).* (Williams & Wilkins Co, 1946). doi:10.1037/11329-000

6. Cutler, D. M., Lleras-Muney, A., Cutler, D. & Lleras-Muney, A. Education and Health: Insights from International Comparisons. (2012).

7. Gottfredson, L. S. & Deary, I. J. *Intelligence Predicts Health and Longevity, but Why? CURRENT DIRECTIONS IN PSYCHOLOGICAL SCIENCE* **13**, (2004).

8. Arendt & Nielsen, J. Does education cause better health? A panel data analysis using school reforms for identification. *Econ. Educ. Rev.* **24**, 149–160 (2005).

9. Panizzon, M. S. *et al.* Genetic and Environmental Influences of General Cognitive Ability: Is g a valid latent construct? *Intelligence* **43**, 65 (2014).

10. Bouchard, T. & McGue, M. Familial studies of intelligence: a review. *Science (80-. ).* **212**, 1055–1059 (1981).

11. Plomin, R. & Spinath, F. M. Intelligence: Genetics, Genes, and Genomics. *J. Pers. Soc. Psychol.* **86**, 112–129 (2004).

12. Colom, R. *et al.* Neuroanatomic overlap between intelligence and cognitive factors: Morphometry methods provide support for the key role of the frontal lobes. *Neuroimage* **72**, 143–152 (2013).

13. Colom, R. *et al.* Gray matter correlates of fluid, crystallized, and spatial intelligence: Testing the P-FIT model. *Intelligence* **37**, 124–135 (2009).

14. Burgaleta, M., Johnson, W., Waber, D. P., Colom, R. & Karama, S. Cognitive ability changes and dynamics of cortical thickness development in healthy children and adolescents. *Neuroimage* **84**, 810–819 (2014).

15. Martínez, K. *et al.* Reproducibility of brain-cognition relationships using three cortical surface-based protocols: An exhaustive analysis based on cortical thickness. *Hum. Brain Mapp.* **36**, 3227–3245 (2015).

16. Sternberg, R. & Grigorenko, E. *The general factor of intelligence*. (Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers, 2002).

17. Van Der Maas, H. L. J. *et al.* A Dynamical Model of General Intelligence: The Positive Manifold of Intelligence by Mutualism. (2006). doi:10.1037/0033-295X.113.4.842

18. Van Der Maas, H., Kan, K.-J., Marsman, M. & Stevenson, C. E. Network Models for Cognitive Development and Intelligence. *J. Intell.* **5**, 16 (2017).

19. Fjell, A. M. *et al.* Multimodal imaging of the self-regulating developing brain. *Proc. Natl. Acad. Sci. U. S. A.* **109**, 19620–19625 (2012).

20. Curley, L. B. *et al.* Cortical morphology of the pars opercularis and its relationship to motor-inhibitory performance in a longitudinal, developing cohort. *Brain Struct. Funct.* **223**, 211–220 (2018).

21. Newman, E. *et al.* Anxiety is related to indices of cortical maturation in typically developing children and adolescents. *Brain Struct. Funct.* **221**, 3013–3025 (2016).

22. Newman, E. *et al.* Go/No Go task performance predicts cortical thickness in the caudal inferior frontal gyrus in young adults with and without ADHD. *Brain Imaging Behav.* **10**, 880–892 (2016).

23. Vuoksimaa, E. *et al.* Is bigger always better? The importance of cortical configuration with respect to cognitive ability. *Neuroimage* **129**, 356–366 (2016).

24. Palmer, C. E. *et al.* Determining the association between cortical morphology and cognition in 10,145 children from the Adolescent Brain and Cognitive Development (ABCD) study using the MOSTest. *bioRxiv* 816025 (2019). doi:10.1101/816025

25. Zhao, W. *et al.* The Bayesian polyvertex score (PVS-B): a whole-brain phenotypic prediction framework for neuroimaging studies. *bioRxiv* 813915 (2019). doi:10.1101/813915

26. Meer, D. van der *et al.* Making the MOSTest of imaging genetics. *bioRxiv* 767905 (2019). doi:10.1101/767905

27. Hickok, G. & Poeppel, D. The cortical organization of speech processing. *Nat. Rev. Neurosci.* **8**, 393–402 (2007).

28. Martin, A., Schurz, M., Kronbichler, M. & Richlan, F. Reading in the brain of children and adults: a meta-analysis of 40 functional magnetic resonance imaging studies. *Hum. Brain Mapp.* **36**, 1963–1981 (2015).

29. Brown, W. E. *et al.* Preliminary evidence of widespread morphological variations of the brain in dyslexia. *Neurology* **56**, 781–3 (2001).

30. Botvinick, M. M., Cohen, J. D. & Carter, C. S. Conflict monitoring and anterior cingulate cortex: an update. *Trends Cogn. Sci.* **8**, 539–546 (2004).

31. Menon, V. & Uddin, L. Q. Saliency, switching, attention and control: a network model of insula function. *Brain Struct. Funct.* **214**, 655–667 (2010).

32. Taylor, K. S., Seminowicz, D. A. & Davis, K. D. Two systems of resting state connectivity between the insula and cingulate cortex. *Hum. Brain Mapp.* **30**, 2731–2745 (2009).

33. Thompson, W. K. *et al.* The structure of cognition in 9 and 10 year-old children and associations with problem behaviors: Findings from the ABCD study's baseline neurocognitive battery. *Dev. Cogn. Neurosci.* **36**, (2019).

34. Bartholomew, D. J., Deary, I. J. & Lawn, M. A new lease of life for Thomson's bonds model of intelligence. *Psychol. Rev.* **116**, 567–79 (2009).

35. Kovacs, K. & Conway, A. R. A. Process Overlap Theory: A Unified Account of the General Factor of Intelligence. *Psychol. Inq.* **27**, 151–177 (2016).

36. Siegler, R., Jenkins, E. A. & Jenkins, E. A. *How Children Discover New Strategies*. (Routledge, 1989). doi:10.4324/9781315807744

37. Eyler, L. T. *et al.* Genetic and environmental contributions to regional cortical surface area in humans: A magnetic resonance imaging twin study. *Cereb. Cortex* **21**, 2313–2321 (2011).

38. Eyler, L. T. *et al.* A comparison of heritability maps of cortical surface area and thickness and the influence of adjustment for whole brain measures: A magnetic resonance imaging twin study. *Twin Res. Hum. Genet.* **15**, 304–314 (2012).

39. Winkler, A. M. *et al.* Cortical thickness or grey matter volume? The importance of selecting the phenotype for imaging genetics studies. *Neuroimage* **53**, 1135–1146 (2010).

40. O'Leary, D. D. M., Chou, S. J. & Sahara, S. Area patterning of the mammalian cortex. *Neuron* **56**, 252–269 (2007).

41. Rakic, P., Ayoub, A. E., Breunig, J. J. & Dominguez, M. H. Decision by division: making cortical maps. *Trends in Neurosciences* **32**, 291–301 (2009).

42. Dickens, W. T. & Flynn, J. R. Heritability estimates versus large environmental effects: the IQ paradox resolved. *Psychol. Rev.* **108**, 346–69 (2001).

43. Kan, K.-J., Wicherts, J. M., Dolan, C. V. & van der Maas, H. L. J. On the Nature and Nurture of Intelligence and Specific Cognitive Abilities. *Psychol. Sci.* **24**, 2420–2428 (2013).

44. Loughnan, R. J. *et al.* Polygenic Score of Intelligence is More Predictive of Crystallized than Fluid Performance Among Children. *bioRxiv* 637512 (2019). doi:10.1101/637512

45. Harden, K. P., Turkheimer, E. & Loehlin, J. C. Genotype by environment interaction in adolescents' cognitive aptitude. *Behav. Genet.* **37**, 273–283 (2007).

46. Compton, W. M., Dowling, G. J. & Garavan, H. Ensuring the Best Use of Data. *JAMA Pediatr.* (2019). doi:10.1001/jamapediatrics.2019.2081

47. Garavan, H. *et al.* Recruiting the ABCD sample: Design considerations and procedures. *Dev. Cogn. Neurosci.* **32**, 16–22 (2018).

48. Eriksen, B. A. & Eriksen, C. W. Effects of noise letters upon the identification of a target letter in a nonsearch task. *Percept. Psychophys.* **16**, 143–149 (1974).

49. Heaton, R. K. *et al.* Reliability and validity of composite scores from the NIH Toolbox Cognition Battery in adults. *J. Int. Neuropsychol. Soc.* **20**, 588–98 (2014).

50. Akshoomoff, N. *et al.* VIII. NIH Toolbox Cognition Battery (CB): composite scores of crystallized, fluid, and overall cognition. *Monogr. Soc. Res. Child Dev.* **78**, 119–132 (2013).

51. Daniel, M. H., Wahlstrom, D. & Zhang, O. *Equivalence of Q-interactive™ and Paper Administrations of Cognitive Tasks: WISC ®-V Q-interactive Technical Report 8*. (2014).

52. Acker, W. *A computerized approach to psychological screening—The Bexley-Maudsley Automated Psychological Screening and The Bexley-Maudsley Category Sorting Test. International Journal of Human-computer Studies / International Journal of Man-machine Studies - IJMMS* **17**, (1982).

53. Weschler, D. *Weschler Intelligence Scale for Children, 5th ed*. (Pearson, Bloomington, MN, 2014).

54. Tisdall, M. D. *et al.* Volumetric navigators for prospective motion correction and selective reacquisition in neuroanatomical MRI. *Magn. Reson. Med.* **68**, 389–399 (2012).

55. White, N. *et al.* PROMO: Real-time prospective motion correction in MRI using image-based tracking. *Magn. Reson. Med.* **63**, 91–105 (2010).

56. Casey, B. J. *et al.* The Adolescent Brain Cognitive Development (ABCD) study: Imaging acquisition across 21 sites. *Dev. Cogn. Neurosci.* **32**, 43–54 (2018).

57. Hagler, D. J. *et al.* Image processing and analysis methods for the Adolescent Brain

Cognitive Development Study. *Neuroimage* 116091 (2019). doi:10.1016/J.NEUROIMAGE.2019.116091

58. Fischl, B., Sereno, M. I. & Dale, A. M. *Cortical Surface-Based Analysis II: Inflation, Flattening, and a Surface-Based Coordinate System*. (1999).

59. Dale, A. M., Fischl, B. & Sereno, M. I. Cortical Surface-Based Analysis. *Neuroimage* **9**, 179–194 (1999).

60. Fischl, B. & Dale, A. M. Measuring the thickness of the human cerebral cortex from magnetic resonance images. *Proc. Natl. Acad. Sci. U. S. A.* **97**, 11050–11055 (2000).

61. Fischl, B. *et al.* Sequence-independent segmentation of magnetic resonance images. *Neuroimage* **23**, S69--S84 (2004).

62. Jovicich, J. *et al.* Reliability in multi-site structural MRI studies: Effects of gradient non-linearity correction on phantom and human data. *Neuroimage* **30**, 436–443 (2006).

63. Vilhjálmsson, B. J. *et al.* Modeling Linkage Disequilibrium Increases Accuracy of Polygenic Risk Scores. *Am. J. Hum. Genet.* **97**, 576–592 (2015).

# FIGURES

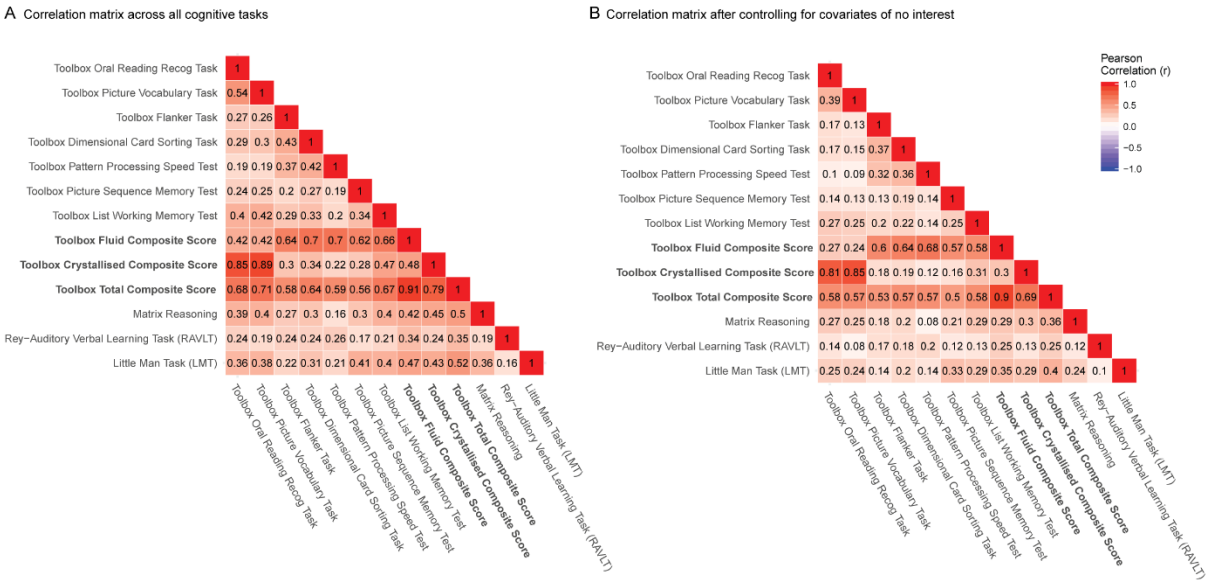


**Figure 1. Phenotypic correlations between all of the cognitive measures analysed.** A) Pearson correlation coefficients for all cognitive tasks. B) Correlations after controlling for covariates of no interest (age, sex, race/ethnicity, household income, parental education, scanner). Cognitive performance was pre-residualised for all covariates and then correlated using a Pearson correlation. Darker colours indicate a higher correlation coefficient. The composite scores are highlighted in bold. The other cognitive measures are single tasks. Correlation coefficients among the single task measures decreased after controlling for the covariates of no interest demonstrating that these additional measures explained a proportion of the variance in cognitive performance among these tasks.

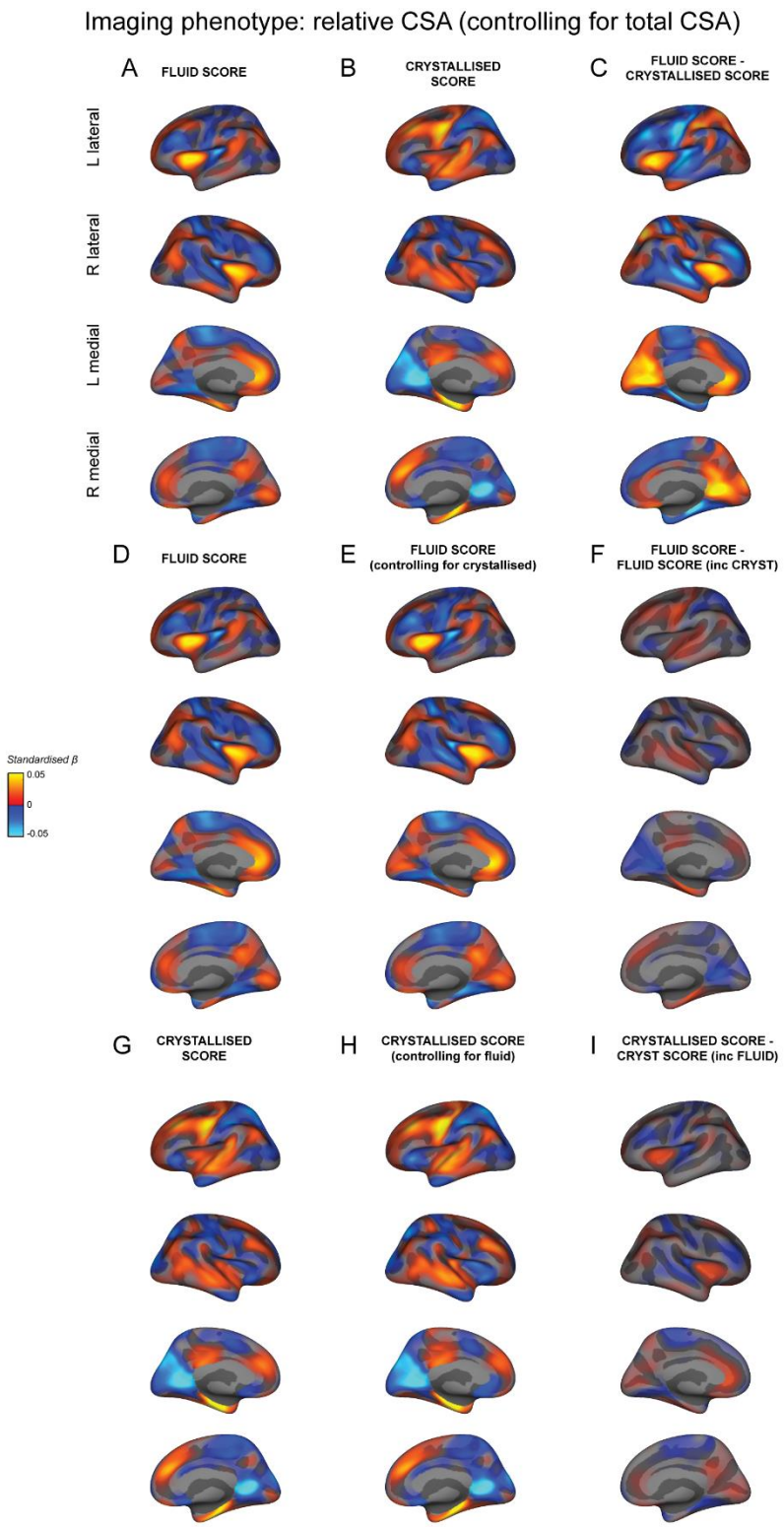


**Figure 2. Distinct estimated effect size maps of association between the regionalisation of CSA and fluid and crystallised composite scores.** All maps display the vertex-wise mass univariate standardised beta coefficients unthresholded for each association. A+D) The association between relative CSA and the fluid composite score and (B+G) the crystallised composite score. C) The difference in standardised beta coefficients between A+B. E) The association between relative CSA and the fluid composite score controlling for the crystallised composite score. F) The difference in standardised beta coefficients between D+E. H) The association between relative CSA and the crystallised composite score controlling for the fluid composite score. I) The difference in standardised beta coefficients between G+H. The estimated effect size maps showing the association between the regionalisation of CSA and the fluid and crystallised composite scores have distinct patterns. These behavioural measures show very little overlapping variance with the regionalisation of CSA.

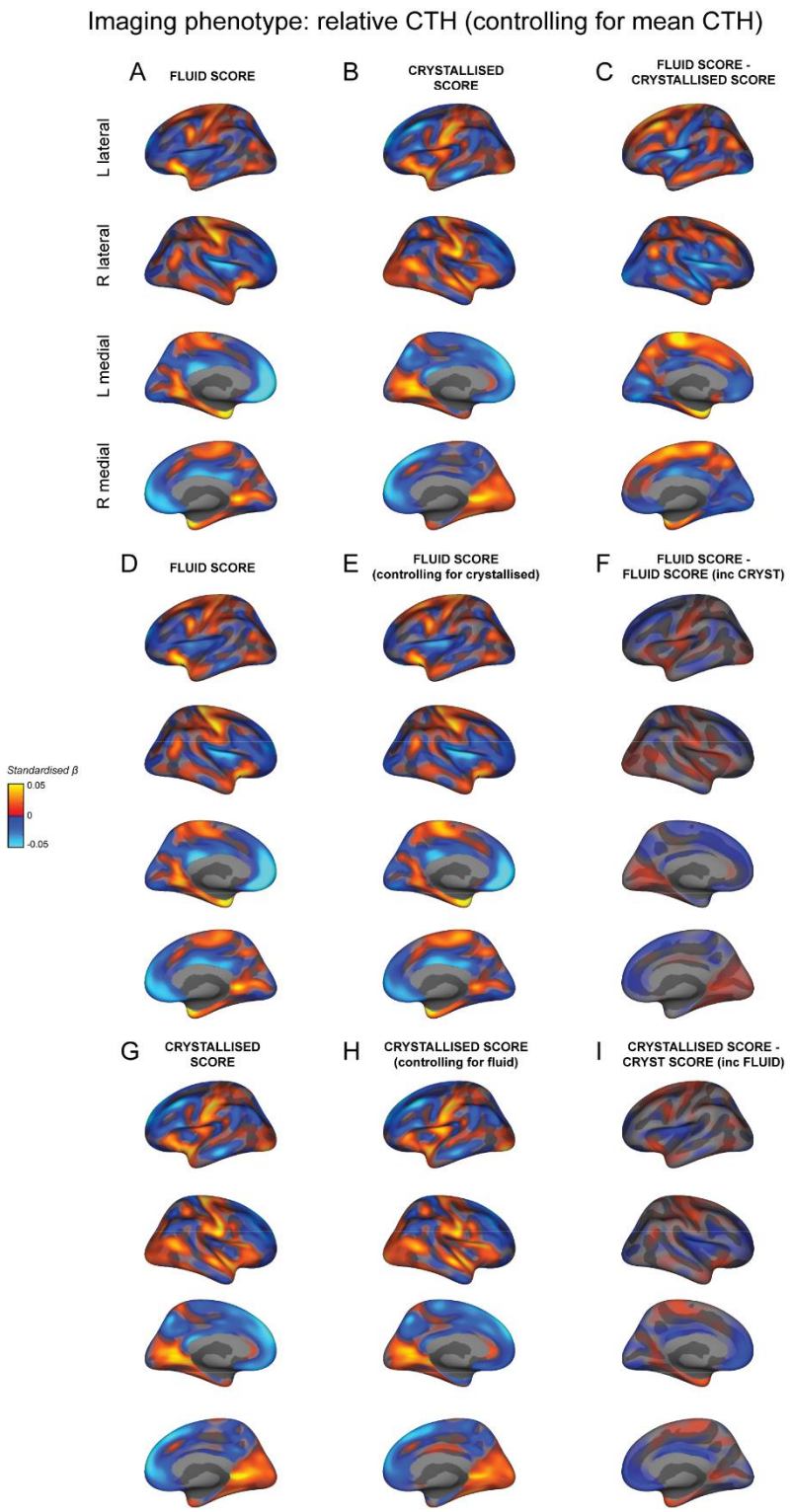


**Figure 3. Distinct estimated effect size maps of association between the regionalisation of CTH and fluid and crystallised composite scores.** All maps display the vertexwise mass univariate standardised beta coefficients unthresholded for each assocition. A+D) The association between relative CTH and the fluid composite score and (B+G) the crystallised composite score. C) The difference in standardised beta coefficients between A+B. E) The association between relative CTH and the fluid composite score controlling for the crystallised composite score. F) The difference in standardised beta coefficients between D+E. H) The association between relative CTH and the crystallised composite score controlling for fluid composite score. I) The difference in standardised beta coefficients between G+H. The estimated effect size maps showing the association between the regionalisation of CTH and the fluid and crystallised composite scores have distinct patterns. These behavioural measures show very little overlapping variance with the regionalisation of CTH.

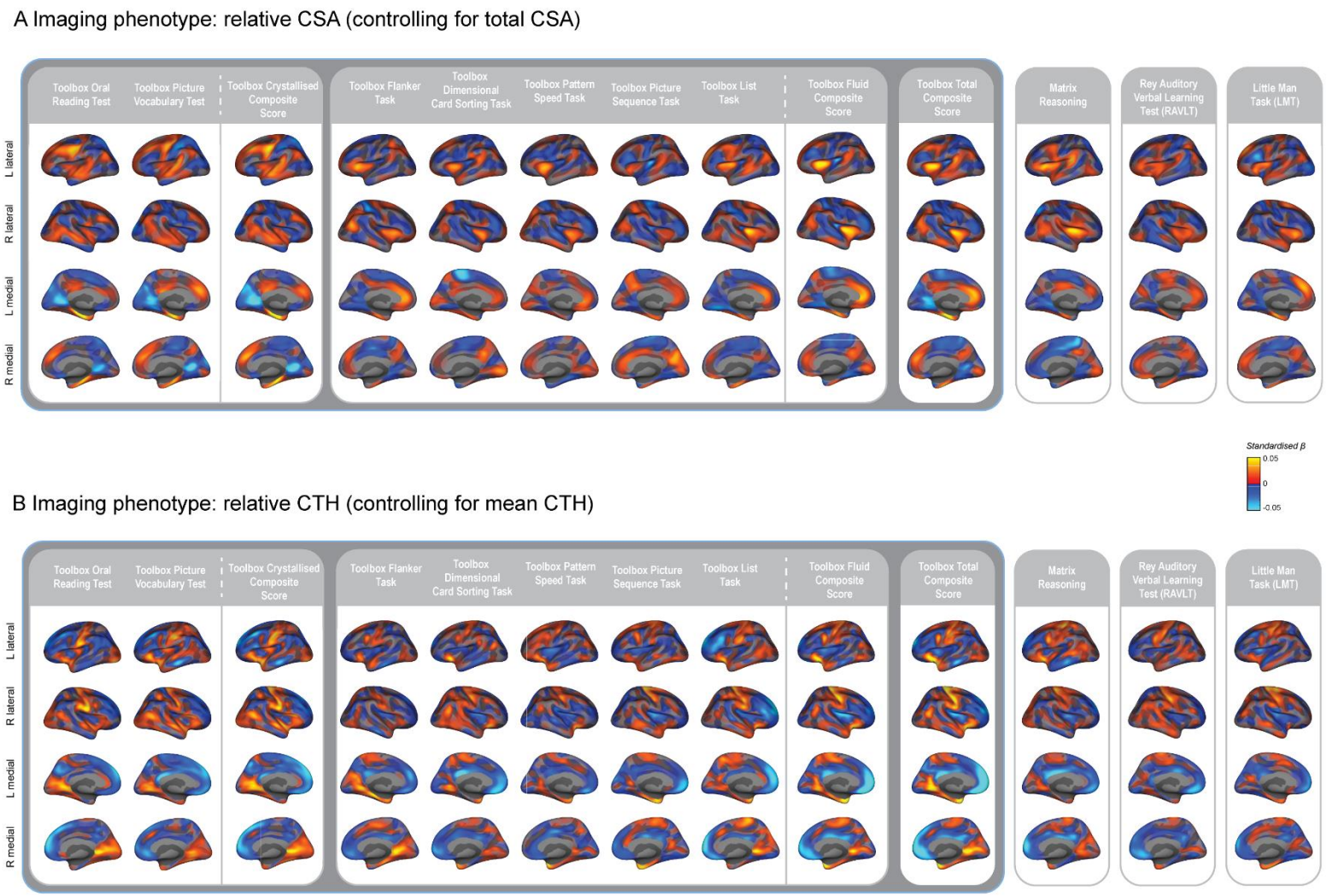


**Figure 4.** Estimated effect size maps showing the mass univariate standardised beta coefficients for the association between each cognitive task and (A) the regionalisation of CSA and (B) the regionalisation of CTH. Performance on all of the cognitive tasks and the composite scores showed significant associations with the regionalisation of cortical morphology across the whole cortical surface according to the MOSTest as determined in Palmer et al (2019).

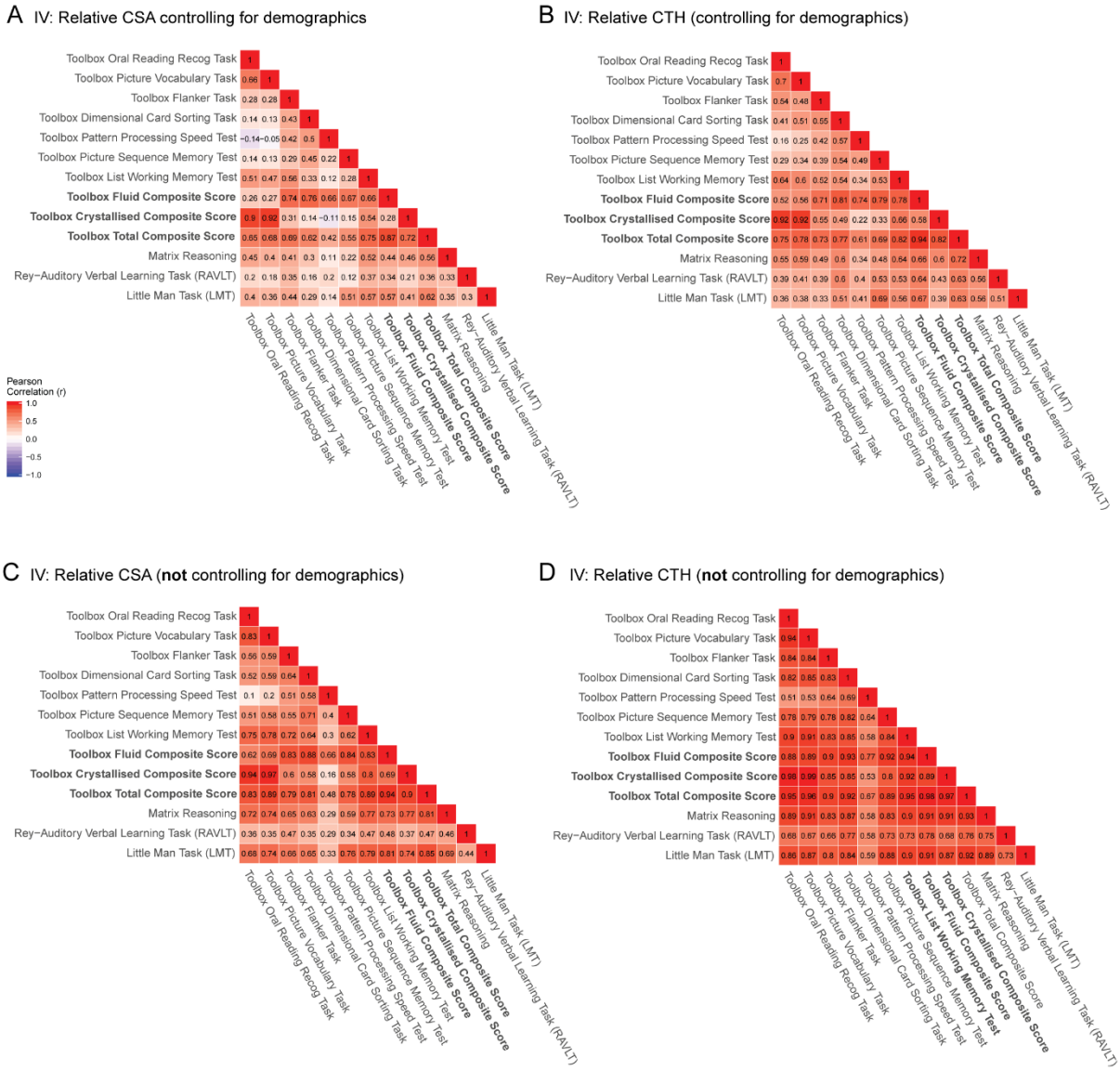


**Figure 5. Similarity of regionalisation association patterns across cognitive tasks is modulated by demographic factors.** Pairwise vertex-wise correlations between the estimated beta coefficients across all cognitive tasks predicted by A) the regionalisation of CSA and B) the regionalisation of CTH controlling for the demographic variables of race/ethnicity, household income and parental education. C&D) Pairwise vertex-wise correlations for all of the cognitive tasks and C) the regionalisation of CSA and D) the regionalisation of CTH <u>not</u> controlling for the specified demographic variables. The correlation amongst associations was much larger when these demographic variables were not included in the GLMs used to produce the estimated effect size maps.

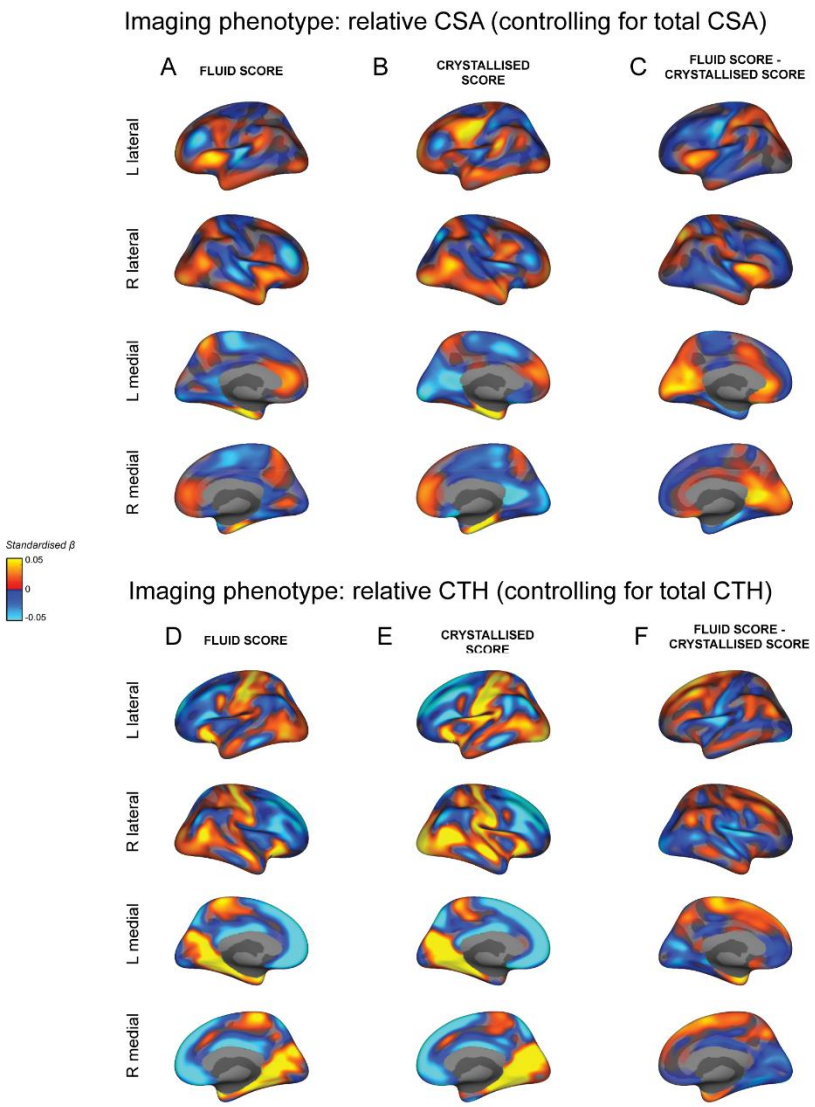


**Figure 6. Similar estimated effect size maps for the association between the regionalisation of cortical morphology and the fluid and crystallised composite scores when <u>not</u> controlling for demographic factors.** All maps display the vertex-wise mass univariate standardised beta coefficients unthresholded for each association. The top part of the figure shows the association between the regionalisation of CSA and the fluid composite score (A) and the crystallised composite score (B). C) The difference in standardised beta coefficients between A+B. The bottom part of the figure shows the association between the regionalisation of CTH and the fluid composite score (D) and the crystallised composite score (E). F) The difference in standardised beta coefficients between D+E. The demographic variables of race/ethnicity, household income and parental education were not included in the mass univariate GLMs to produce these estimated effect size maps. The exclusion of these confounding factors increased the magnitude of the estimated beta coefficients and the similarity between the maps for the fluid and crystallised composite scores. However, the difference maps (C+F) closely reflect the pattern of differences seen in figure 2C and figure 3C respectively.

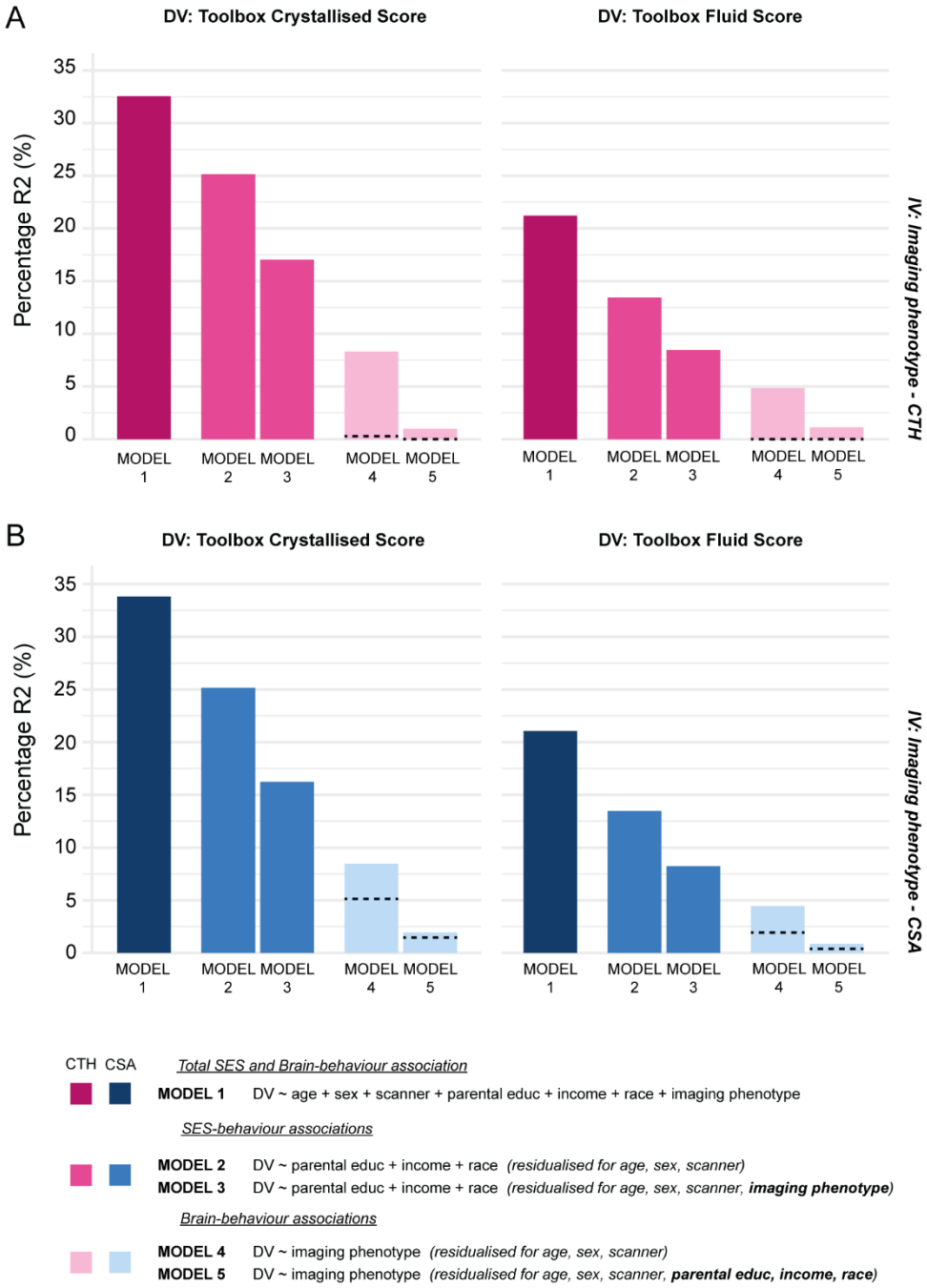


**Figure 7. Controlling for sociodemographic factors reduces the variance in cognitive performance predicted by cortical morphology.** These bar charts show the percentage $R^2$ in the crystallised (left column) and fluid (right column) composite scores predicted by several models including sociodemographic and brain measures. Model 1 shows the total variance in behaviour explained by the full model including the covariates of no interest (age, sex and scanner), the sociodemographic variables (parental education, household income and race/ethnicity) and either the regionalisation of CTH and mean CTH (A; pink) or the regionalisation of CSA and total CSA (B; blue). Model 2 shows the *unique* variance in behaviour explained by the sociodemographic factors after pre-residualising both the dependent (DV) and independent (IV) variables for the covariates of no interest. Additionally, pre-residualising for the imaging phenotypes (model 3) lead to a reduction in this $R^2$. Conversely, model 4 shows the *unique* variance in behaviour explained by the structural measures (global below the dotted line; and regionalisation above the dotted line) after pre-residualising for the covariates of no interest only. Additionally, pre-residualising for the sociodemographic factors (model 5) lead to a large decrease in the variance explained by these structural measures, again showing the shared variance between these measures. Controlling for the sociodemographic factors of race/ethnicity, household income and parental education reduced the variance explained in the composite scores by ~4-5 fold for CSA and ~4-8 fold for CTH.