BRAIN
ORIGINAL ARTICLE

Graph lesion-deficit mapping of fluid intelligence

Lisa Cipolotti,[1,2] James K. Ruffle,[2,3] Joe Mole,[1,2] Tianbo Xu,[2] Harpreet Hyare,[2,3] Tim Shallice,[4,5] Edgar Chan[1,2] and Parashkev Nachev[2]

Fluid intelligence is arguably the defining feature of human cognition. Yet the nature of its relationship with the brain remains a contentious topic. Influential proposals drawing primarily on functional imaging data have implicated 'multiple demand' frontoparietal and more widely distributed cortical networks, but extant lesion-deficit studies with greater causal power are almost all small, methodologically constrained, and inconclusive. The task demands large samples of patients, comprehensive investigation of performance, fine-grained anatomical mapping, and robust lesion-deficit inference, yet to be brought to bear on it.

We assessed 165 healthy controls and 227 frontal or non-frontal patients with unilateral brain lesions on the best-established test of fluid intelligence, Raven's Advanced Progressive Matrices, employing an array of lesion-deficit inferential models responsive to the potentially distributed nature of fluid intelligence. Non-parametric Bayesian stochastic block models were used to reveal the community structure of lesion deficit networks, disentangling functional from confounding pathological distributed effects.

Impaired performance was confined to patients with frontal lesions [$F(2,387) = 18.491$; $P < 0.001$; frontal worse than non-frontal and healthy participants $P < 0.01$, $P < 0.001$], more marked on the right than left [$F(4,385) = 12.237$; $P < 0.001$; right worse than left and healthy participants $P < 0.01$, $P < 0.001$]. Patients with non-frontal lesions were indistinguishable from controls and showed no modulation by laterality. Neither the presence nor the extent of multiple demand network involvement affected performance. Both conventional network-based statistics and non-parametric Bayesian stochastic block modelling heavily implicated the right frontal lobe. Crucially, this localization was confirmed on explicitly disentangling functional from pathology-driven effects within a layered stochastic block model, prominently highlighting a right frontal network involving middle and inferior frontal gyrus, pre- and post-central gyri, with a weak contribution from right superior parietal lobule. Similar results were obtained with standard lesion-deficit analyses.

Our study represents the first large-scale investigation of the distributed neural substrates of fluid intelligence in the focally injured brain. Combining novel graph-based lesion-deficit mapping with detailed investigation of cognitive performance in a large sample of patients provides crucial information about the neural basis of intelligence. Our findings indicate that a set of predominantly right frontal regions, rather than a more widely distributed network, is critical to the high-level functions involved in fluid intelligence. Further they suggest that Raven's Advanced Progressive Matrices is a useful clinical index of fluid intelligence and a sensitive marker of right frontal lobe dysfunction.

1   Department of Neuropsychology, National Hospital for Neurology and Neurosurgery, London WC1N 3BG, UK
2   Institute of Neurology, University College London, London WC1N 3BG, UK
3   Department of Radiology, University College London Hospitals NHS Foundation Trust, London NW1 2PG, UK
4   Institute of Cognitive Neuroscience, University College London, London WC1N 3AZ, UK
5   Cognitive Neuropsychology and Neuroimaging Lab, International School for Advanced Studies (SISSA-ISAS), 34136 Trieste, Italy

Correspondence to: Prof. Lisa Cipolotti
Department of Neuropsychology
National Hospital for Neurology and Neurosurgery
Queen Square, London WC1N 3BG, UK
E-mail: l.cipolotti@ucl.ac.uk

Introduction

Fluid intelligence refers to the ability to solve challenging novel problems when prior learning or accumulated experience are of limited use.[1] Fluid intelligence ranks amongst the most important features of cognition, correlates with many cognitive abilities (e.g. memory),[2] and predicts educational and professional success,[3] social mobility,[4] health[5] and longevity.[6] It is thought to be a key mental capacity involved in 'active thinking',[7] fluid intelligence declines dramatically in various types of dementia[8] and reflects the degree of executive impairment in older patients with frontal involvement.[9] Despite the importance of fluid intelligence in defining human behaviour, it remains contentious whether this is a single or a cluster of cognitive abilities and the nature of its relationship with the brain.[10]

Fluid intelligence is traditionally measured with tests of novel problem-solving with non-verbal material that minimize dependence on prior knowledge. Such tests are known to have strong fluid intelligence correlations in large-scale factor analyses.[11,12] Raven's Advanced Progressive Matrices[13] (APM), a test widely adopted in clinical practice and research,[14] contains multiple choice visual analogy problems of increasing difficulty. Each problem presents an incomplete matrix of geometric figures with a multiple choice of options for the missing figure. Less commonly, verbal tests of fluid intelligence such as Part 1 of the Alice Heim 4 (AH4-1)[15] are adopted. The Wechsler Adult Intelligence Scale (WAIS)[16] has also been used to estimate fluid intelligence by averaging performance on a diverse range of subtests. However, several subtests (e.g. vocabulary) emphasize knowledge, disproportionately weighting measures of 'crystallized' intelligence,[17,18] whilst others (e.g. picture completion) have rather low fluid intelligence correlations.[19] Hence, it has been argued that tests such as the APM are the most suitable for a theoretically-based investigation of changes in fluid intelligence after brain injury.[20,21]

Proposals regarding the neural substrates of fluid intelligence have suggested close links with frontal and parietal functions. For example, Duncan and colleagues[22] have argued that a network of mainly frontal and parietal areas, termed the 'multiple-demand network' (MD), is 'the seat' of fluid intelligence. The highly influential parieto-frontal integration theory (P-FIT), based largely on neuroimaging studies of healthy subjects, posits that structural symbolism and abstraction emerge from sensory inputs to parietal cortex, with hypothesis generation and problem solving arising from interactions with frontal cortex. Once the best solution is identified, the anterior cingulate is engaged in response selection and inhibition of alternatives.[23,24] Despite its name, P-FIT also posits occipital and temporal involvement, implying widely distributed substrates of fluid intelligence.[25] A modification to P-FIT proposes a closer connection between frontal than parietal, regions and fluid intelligence-related processes,[26,27] with the frontal lobes mediating high fluid intelligence 'domain-independent' executive processes whilst posterior areas, including the parietal lobes, mediating low fluid intelligence 'domain-dependent' processing of spatial, object, or verbal information.

A meta-analysis of the functional imaging literature has implicated a network of modality-independent regions involving the inferior and middle frontal and inferior parietal lobes, with additional frontal eye field activation in non-verbal tasks, and anterior cingulate and left inferior frontal activation in verbal tasks.[28] This frontoparietal attention network[29] requires expansion to account for the separate neuronal substrate underpinning visuospatial/verbal analytical reasoning.[30,31]

An important caveat of the functional imaging findings is that they do not imply causal efficacy.[32] For example, though neuropsychological data commonly lateralize language to the left hemisphere,[33] neuroimaging activation is often bilateral.[34] So, merely considering the presence or absence of activation may hide lateralized functions. Hence, lesion studies offer an advantage in furthering our understanding of the neurocognitive architecture underpinning fluid intelligence. So far, these studies have been surprisingly sparse.

Lesion studies investigating performance on fluid intelligence tasks have mainly enrolled veterans with penetrating head injury.[35–43] For example, Weinstein and Teuber[43] reported that veterans with left temporo-parietal entrance wounds suffered a decline in Army General Classification Test scores. Barbey and colleagues[44] investigated Veterans' performance on three subtests from the WAIS (Matrix Reasoning, Block Design and Picture Completion). The authors reported that performance was associated with damage to the superior longitudinal/arcuate fasciculus. However, several of the tests adopted are not considered fluid intelligence measures,[45] and the lesion characterization was rather basic and lacked modelling of the diffuse axonal injury the traumatic aetiology implies.[46] In the most recent of these studies, impaired WAIS performance, not a specific test of fluid intelligence, was associated with damage to left fronto-parietal regions and white matter association tracts.[47] Glascher and colleagues[48] reported that the left frontal pole was associated with performance on general intelligence (g) in a large sample of patients with stroke, encephalitis, temporal lobectomy and traumatic brain injury (TBI). Similarly, Browen and colleagues[49] investigated performance in general intelligence using the WAIS in a sample of patients with similar pathologies, and reported an association with white matter tracts deep to the left temporo-parietal junction, including the arcuate fasciculus.

Studies investigating patients with lesions caused by brain tumours or stroke have generally relied on WAIS as a measure of fluid intelligence, with inconclusive results. Some studies have associated WAIS non-verbal scale performance with right posterior damage.[50,51] However, Tranel and colleagues[52] reported no significant differences between frontal and non-frontal damage on a non-verbal subtest of the WAIS analogous to the APM (Matrix Reasoning). Preserved performance on the WAIS has been documented in frontal patients.[53,54] In contrast, the very few studies adopting tasks loading more heavily on fluid intelligence have reported frontal deficits. Duncan and colleagues[20] documented a substantial discrepancy between scores on Scale 2 of Cattell's Culture Fair and the WAIS in three frontal but not in five non-frontal patients. However, the very small sample limited generalizability and prevented investigation of laterality effects. In a recent study we documented lateralised frontal effects on APM and AH4-1.[55] Compared with healthy participants, only right

frontal damage significantly impaired APM performance, and only left frontal damage impaired AH4-1 performance. The relatively small sample prevented investigation of finer anatomical effects.

Lesion studies investigating the underlying behavioural and anatomical aspects of the widely used APM are old, inconclusive, and lacking in anatomical analysis. Results have variously shown no difference between right or left hemisphere patients[56–61] or impairment in left hemisphere patients[56,62,63] or right hemisphere patients.[56–65] Villardita[66] reported no difference in the performance of left or right hemisphere patients on the Coloured Progressive Matrices version of APM. However, on Set I, involving visuoperceptual factors, right performed worse than left hemisphere patients. There is similar uncertainty about the influence of aphasia, thought by some to degrade performance,[58–67] but not by others.[60–69] Large samples of patients with focal unilateral lesions, thorough investigation of performance on the APM, fine-grained anatomical mapping, and robust lesion-deficit inference are vital for definitive scientific conclusions.

Here we assessed the largest number of patients yet reported with focal, unilateral, right or left, frontal or non-frontal lesions ($n = 227$; 146 frontal, 81 non-frontal) and 165 healthy participants on APM. We investigated overall performance, item difficulty, and relation to MD involvement. Building upon our novel multimodal methodology,[70] we employed an array of lesion-deficit models responsive to the potentially distributed nature of fluid intelligence. We focused on modelling the anatomy of neural dependence as a graph, where interactions between distributed areas are explicitly tested. This approach permits delineation of distributed substrates. It also distinguishes functionally critical areas from those the distinctive pathological structure of lesions renders spuriously correlated: a problem shown to corrupt lesion-deficit maps based on simple mass univariate methods.[71] In our approach each brain locus—intact or lesioned—is conceived as a node or vertex of a graph, with the relationships between loci—functional or merely lesion-pathology driven—defining its edges. This permits us to model network-dependence explicitly, disentangling functional and pathological effects to reveal the underlying substrate.

Materials and methods
Participants

Data from 332 patients with unilateral, focal lesions who attended the Neuropsychology Department of the National Hospital for Neurology and Neurosurgery was retrospectively screened. Inclusion criteria were: presence of a stroke or tumour; ≥70% of the total lesion, segmented from MRI or CT scans obtained during routine clinical care (see 'Neuroimaging investigations' section), falling within either frontal or non-frontal areas; age between 18–80 years; absence of gross perceptual impairments (no neglect, >5th cut-off on the Incomplete Letters test),[72] language impairments (>5th %ile on the Graded Naming Test, GNT[73]; psychiatric disorders, history of alcohol or substance abuse, or other neurological disorders; and native English language proficiency. Age at assessment, gender, and years of education were recoded.

Application of these criteria yielded 227 patients, 146 frontal [left frontal (LF) 69; right frontal (RF) 77], and 81 non-frontal [left non-frontal 39 (LNF); right non-frontal 421 (RNF); see Table 1]. There was no significant difference between tumour and stroke patients for mean time between injury and neuropsychological assessment ($P = 0.12$; Table 1). One hundred and sixty-five healthy control participants, with no neurological or psychiatric history, were recruited to match patients as closely as possible for age, gender, years of education and National Adult Reading Test scores (NART).[74]

The study was approved by The National Hospital for Neurology and Neurosurgery and Institute of Neurology Joint Research Ethics Committee and conducted in accordance with the 'Declaration of Helsinki'.

Behavioural investigations

Patients were assessed with tests administered and scored in the published standard manner. Due to the retrospective nature of our study, certain data were unavailable for some participants.

Background tests

Premorbid optimal level of functioning was assessed using the NART, perception and naming using Incomplete Letters and GNT. Two widely used executive tasks, known to require processes distinct from fluid intelligence were also administered.[9–75] Verbal generation was assessed using the phonemic fluency test.[76] The total number of words recalled excluding errors was recorded. Strategy formation/response inhibition was assessed using the Hayling Sentence Completion Test. Suppression errors were calculated.[77]

Table 1 Demographics and cognitive test scores

	n	Healthy control mean	n	Frontal mean	n	Non-frontal mean
Age, years (SD)	165	49.61 (15.19)	146	49.17 (15.76)	81	51.76 (14.93)
Gender, male/female		77/88		81/65		45/36
Aetiology: stroke/tumour/abscess/AVM				38/104/4/0		35/45/0/1
Chronicity, days (SD)			124	238.52 (640.41)	73	267.26 (740.93)
Education, years (SD)	148	14.57 (2.65)	135	14.38 (3.91)	78	14.83 (2.83)
Premorbid NART IQ (SD)	165	107.02 (10.01)	146	105.48 (12.71)	81	108.19 (10.60)
GNT (Correct/30) (SD)	131	20.91 (5.14)	121	19.69 (4.78)	71	20.23 (5.26)
IL, Correct/20 (SD)	123	19.59 (0.59)	111	19.48 (0.74)	68	19.49 (0.76)
S Fluency, overall performance (SD)	69	17.19 (4.79)	114	**11.54[a,***;b,***] (5.71)**	55	15.04 (5.17)
Hayling suppression error scaled scores (SD)	63	5.95 (0.89)	81	**3.91[a,***;b,***] (2.70)**	51	5.84 (2.08)

AVM = arteriovenous malformation; NART = National Adult Reading Test; SD = standard deviation; SS = Scaled Score; GNT = Graded Naming Test; IL = Incomplete Letters. Scores with significant P-values are in bold.
***P < 0.001.
[a]Significant difference between frontals and non-frontals.
[b]Significant difference between frontals and healthy controls.

Fluid intelligence

Fluid intelligence was assessed using APM.[13] We analysed the following:

Overall performance

The total number of correct responses in Set 1 (/12) was recorded and converted into age-adjusted scaled scores based on published norms.

Item difficulty

Based on visual inspection of the percent correct in healthy control performance, we graded the 12 items from easiest to hardest. We then formed three variables: 'easy group', containing the four easiest items (1, 7, 2, 4), 'medium group', containing the next four items (3, 5, 6, 10) and 'hard group', containing the four hardest items (12, 9, 8, 11). We calculated each patient's score for the three variables (0–4). We compared the performance of the LF, RF and Controls on the three groups to investigate differences in performance based on item difficulty.

Multiple-demand network

We compared overall performance in patients with versus without MD damage, controlling for age and NART (see 'Neuroimaging investigations' section). We used both frequentist and Bayesian linear regression to investigate whether extent of MD involvement predicted APM performance, over and above that predicted by age and NART.

Neuroimaging investigations

Imaging data were available for 176 patients ($n = 173$ MRI, $n = 3$ CT; $n = 110$ frontal, $n = 66$ non-frontal). MRI scans were obtained on either a 3 T or 1.5 T Siemen scanners following a diversity of clinically-determined protocols outside our control. CT studies were obtained on Toshiba or Siemens spiral scanners. Note that since the input to the imaging models are not raw image data but comparatively large, manually-traced, binary lesion masks, in keeping with established practice in the field we made the assumption that the effect of variations in acquisition parameters is likely negligible and need not be explicitly modelled. Lesions were traced and independently classified using MIPAV (https://mipav.cit.nih.gov/) by J.M., E.C. and checked by P.N., who was blind to the study results. In tumour patients, the segmented lesion included the surgical cavity. The lesion masks were non-linearly normalized to Montreal Neurological Institute (MNI) stereotaxic space at $2 \times 2 \times 2$ mm resolution using SPM-12 software (http://www.fil.ion.ucl.ac.uk).[78] The lesion distribution is displayed in Fig. 1. Involvement of the MD was established by comparing each patient's normalized lesion mask with a template of MD regions in MNI space kindly provided by Professor Duncan's group.[79] For each patient, we determined whether their lesion involved the MD and calculated extent of MD involvement (i.e. MD lesion volume/total lesion volume $\times$ 100).

Behavioural analysis

All statistical analyses were conducted using SPSS version 25. Neuropsychological data were assessed for skewness and kurtosis and tested for normality using the Shapiro-Wilk test.

One-way univariate analysis of variance (ANOVA), independent samples *t*-tests or chi-square analyses were conducted for continuous and categorical data, respectively to investigate differences between frontal, non-frontal and control participants on age, gender, aetiology, chronicity, lesion volume, years of education, and neuropsychological variables (NART IQ, GNT, Incomplete Letters, S fluency and Hayling suppression errors). Following significant differences, *post hoc* tests with Bonferroni correction (0.05/3 = P = 0.016) compared frontal versus non-frontal, frontal versus control and non-frontal versus control. LF and RF were also compared on all demographic and neuropsychological variables using *t*-tests.

Standard and lateralization analyses were performed on APM overall performance. In the standard analysis we established the sensitivity and specificity of the APM to the frontal lobes. This analysis was critical because, only if there was a significant frontal deficit compared to Controls the subsequent lateralization analysis was carried out to investigate unilateral left and/or right frontal contributions to APM.
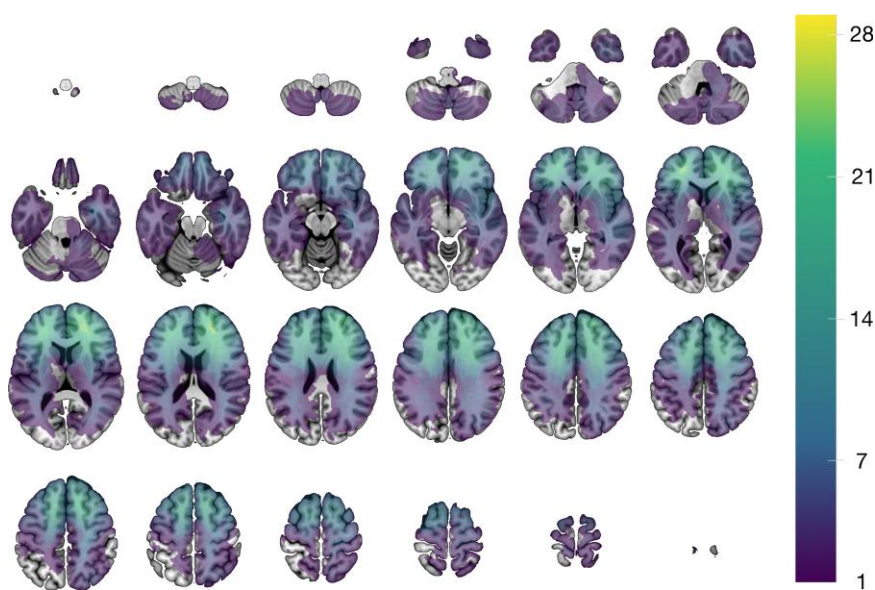

Figure 1 Lesion distribution map. Voxel-wise sum of the 221 modelled lesions overlaid on the SPM152 T$_1$ template distributed with MRIcroGL(https://www.nitrc.org/projects/mricrogl). The images are displayed in neurological convention (left is left).

In the standard analysis, ANCOVA was used to compare frontal versus non-frontal versus control, adjusting for age and NART. Following significant differences, *post hoc* tests with Bonferroni correction ($0.05/3 = P = 0.016$) compared frontal versus non-frontal, frontal versus control and non-frontal versus control. In the laterality analysis, ANCOVA was used to compare LF versus RF versus LNF versus RNF versus Control, adjusting for age and NART. Following significant differences, pairwise comparisons with Bonferroni correction ($0.05/4$ yields $P = 0.0125$) compared each patient group against Control (i.e. LF versus Control, RF versus Control, LNF versus Control, RNF versus Control), LF with RF and LNF with RNF.

To investigate potential differences in performance according to item difficulty in Frontal patients, we used a 3×3 ANCOVA with Difficulty (Easy, Medium, Hard) as the within group factor and Group (LF, RF and Controls) as the between group factor, covaried for age and NART. Significant main effects of Group were followed by simple effects analyses with Bonferroni correction.

To investigate the contribution of the MD to overall performance one-way ANCOVA was used to compare patients with ($n = 153$) versus without ($n = 23$) MD lesions, while adjusting for age and NART. We also performed a multiple linear regression analysis, using the enter method, entering APM performance as the outcome variable and age, NART and extent of MD involvement as predictor variables.

Neuroimaging analysis

Lesion-deficit inference is complicated by the presence of correlations across damaged voxels, not just functionally—arising from a distributed neural substrate—but also pathologically—arising from the structure of the underlying pathological process. Without explicit modelling of regional interactions within high-dimensional models that demand large-scale data, spatial inferences are likely to be unquantifiably distorted. In the absence of an established approach applicable to the comparatively small-scale data regimes inevitable in neuropsychology, we applied multiple inferential methods, focusing on the graph-based approach with the strongest theoretical foundations.

Parcel-based analysis

PLSM ANALYSES were completed using the NiiStat toolbox for Matlab (http://www.nitrc.org/projects/niistat). To increase statistical power the brain was regionally parcellated following the JHU-MNI atlas[80] into 189 regions of interest (ROIs) spanning both grey and white matter. To assure statistical power, only ROIs with damage in ≥ 10 patients were included. Three Freedman-Lane permutations[81] were performed with age, NART and lesion volume always entered as nuisance regressors. Permutation thresholding (5000 permutations) was used to correct for multiple comparisons and control the family-wise error rate. An alpha of 0.05 was the threshold for significance. To investigate the contribution of the MD to APM overall performance, PLSM analyses were repeated with age, NART and proportion of MD involvement entered as nuisance regressors. PLSM analyses were conducted on Hayling suppression errors (scaled scores), with age, NART and lesion volume or age, NART and proportion of MD involvement entered as nuisance regressors.

Bayesian multivariate lesion-deficit modelling of multiple-demand network dependence

To quantify the regional contribution of components of the MD, Bayesian multivariate regression implemented in BayesReg v1.91

was performed with each connected component of the MD map treated as a predictor variable, and age and NART added as nuisance covariates. A selection of shrinkage priors (ridge, lasso, g, horseshoe, horseshoe+) and noise models (Gaussian, Laplace, Student t distribution) were evaluated, choosing g and Student t based on the widely applicable information criterion (WAIC): a standard interpretable metric for Bayesian model comparison.[82] The posterior distributions of the regression coefficients were estimated with Markov chain Monte Carlo sampling over 100 000 samples with a 100 000 burn-in interval and thinning set at 10, reporting the means and standard deviations of the regression coefficients that survive a 95% Bayesian credibility interval. The effective sample size was >97 for all models.

Graph lesion-deficit modelling

Where the neural support of a function is distributed across a set of connected regions, the optimal way of identifying it is through explicit modelling of both anatomical locations and their interactions. Even where the neural support is local, the structure of the lesion pathology used to reveal it need not be, and itself requires modelling of distributed relations. By structure here is meant characteristic spatial patterns of coincident damage dictated by the underlying pathological process, such as the patterns of ischaemic damage the vascular tree enforces in stroke. The difficulty is amplified when both the neural and the pathological are distributed, for the former must then be disentangled from the latter: a problem for which there is no established solution. Here we adopt an approach based on statistical models of graphs. The fundamental idea is to conceive the brain as a densely interconnected graph, where each node is an anatomical location and each edge indexes the extent to which its connected nodes share a set of properties. In the context of lesion-deficit mapping, the properties of interest are the presence of damage, the associated deficit, and nuisance factors that could confound their relations. First, we apply conventional network-based statistics, fitting a general linear model to APM scores, revealing a network of dependence driven jointly by functional anatomy and spatial patterns of damage. Second, we exploit recent developments in Bayesian stochastic block modelling to identify communities of voxels distinctively influenced by fluid intelligence, disentangled from the incidental spatial structure of lesions.

Network-based statistics

The non-linearly registered lesion masks were linearly resampled to a resolution of 12 mm³. This resolution offers considerably finer anatomical detail than published parcellation schemes[83-86] and avoids the potentially biasing effects of their structuring determinants. We provide a bar plot showing the mean node volume (±95% confidence interval) if our approaches compared with commonly used parcellation schemes in Supplementary Fig. 1.

A graph where voxels are the nodes and their adjacent neighbours the edges was created as an adjacency matrix, labelling any edge that linked two lesioned nodes as 1 and all others as 0. This process yielded a graph of order and size 1017 nodes and 516 636 edges for each patient ($n = 172$). The choice of a 12 mm³ voxel size was constrained by the tractability of the statistical model, in line with the practice of others in related domains.[23-84]

We proceeded to model lesion adjacency matrices with the Network-Based Statistics (NBS) connectome toolbox (v1.2).[87] NBS is an established statistical framework for network analysis, described in extensive detail elsewhere.[88,89] In brief, it implements a non-parametric approach to mass-univariate statistical inference

on the edges of large graphs, yielding family-wise error (FWER)-corrected *P*-values for each edge via permutation testing.[90] The approach can be viewed as the graph analogue of the mass-univariate voxel-wise methods familiar from functional imaging and voxel-based morphometry. It has been widely applied to investigate the organization of brain networks.[87–91]

Here the inputs were the lesion graphs of each patient, with APM as the predictor and NART and age as nuisance covariates. The model was fitted with 50 000 permutations, with a criterion for statistical significance set at family-wise error rate corrected *P* < 0.05, yielding an inferred group-level network significantly associated with fluid intelligence. We evaluated the community structure of this inferred network—the presence of clusters of voxels defined by similar inferred connectivity—with a Bayesian, weighted, non-parametric, hierarchical, generative stochastic block model,[92,93] with additional simulated annealing to approximate the global optimum of the function (see below). Edges were weighted by the significant *t*-statistic adjacency matrix from the NBS model. To examine the potential influence of aetiology, we compared this NBS model to another identically configured except for the addition of aetiology as a nuisance covariate (Supplementary material).

To illustrate the relation between the inferred network and fluid intelligence, we created a set of Bayesian regression models with a target of APM adjusted for NART, and predictors constructed—across separate models—from the dichotomized overlap between a lesion and the NBS-identified network, or from the number of nodes of each patient's graph included in the NBS network. We also ran a multivariate regression model with the lesion adjacency matrix as columns of predictors. We evaluated all models with various prior shrinkage schemes, using the WAIC to select the most appropriate prior distributions and model goodness-of-fit.[82,94,95] All regression models were implemented in BayesReg v1.2, and employed a burn-in of 50 000, taking 100 000 samples from the posterior distribution within a single MCMC chain. Note these analyses are not independent and are designed to be merely illustrative of the NBS model from which they are derived.

Bayesian hierarchical stochastic block modelling

The foregoing simple network-based statistical model is potentially confounded by the anatomical structure of pathological damage. It is also tractable only at relatively coarse spatial resolutions. To overcome these defects, we exploited a recent innovation in the statistical modelling of graphs: Bayesian stochastic block models.[96] These are non-parametric probabilistic statistical models of the network structure of graphs that enable robust inference to distinct patterns of connectivity arising as network 'blocks' or 'communities' within them. In the context of a graph model of the lesioned brain, such communities may be shaped by the neural substrate of the behaviour under study, the anatomical patterns of damage, or an interaction between the two. The approach allows us to disentangle these two distinct types of node connectivity, in our case isolating the neural dependents of APM performance from the incidental structure of the lesions used to reveal them. We employ a specific kind of Bayesian stochastic block model designed to incorporate layered, multiple attribute properties.[96] The layered formulation enables robust inference to the separability of the two types of connectivity by Bayesian model comparison of variants whose layered structure either respects or ignores them.[92] Though a comparatively recent innovation, such models rely on well-established principles of Bayesian inference and graph theory, and are underwritten by their theoretically proven validity[92–99].

Graph theory provides a powerful method of modelling complex systems that combines flexibility with intelligibility.[89] It treats individual factors of interest as the 'nodes' of a network, and their interactions as the connections, or 'edges', between them. In the context of lesion deficit inference, the nodes identify anatomical locations in brain, and the edges describe their pairwise relations. Two nodes may be related by their association with a deficit when lesioned, or by their tendency to be involved in the same lesion, regardless of the deficit. The former is the effect of interest, the latter is a potential confounder we wish to eliminate. To disentangle the two forms of relation we create a layered, weighted, undirected graph whose layers correspond to the two different kinds of association. Confining each form of relation to its own layer compels the model to disentangle them in inferring the community structure of the graph. We can compare a layered model of this kind to a null model where the edges are randomized across layers, employing Bayesian model comparison based on the minimum description length of the model. Finding the layered model superior to the null is evidence of the successful separation of the structuring effects of APM and lesion co-occurrence we seek here. A detailed exposition of the inferential approach is given in the Supplementary material.

To model our data, each non-linearly registered lesion was resampled to 4 mm³ resolution, and the lesion adjacency matrix constructed for each patient as before. This resolution is much finer for than conventional parcellation schemes published in the wider literature schemes (Supplementary Fig. 1).[83–86] We then constructed an undirected, weighted graph combining all individual lesion networks across all patients. This network comprised nodes corresponding to all voxels of the brain, and edges between voxels adjacent. These edges were weighted by two variables: the count of the number of times a voxel and an adjacent neighbour were damaged together —a lesion co-occurrence weight—and the inverse of the patient's APM score divided by NART—an adjusted APM weight. Naturally, the graph was undirected, as the direction of any relationship between collaterally lesioned areas is not informed by the data at hand.

We filtered edges to limit analysis to the top 50% connected nodes, removing edges with fewer than ~3 connections, where sampling was too low to support robust inference, but permitting still full brain coverage. This yielded a graph of order and size 27 509 nodes and 285 545 edges. There were no node self-loops. We rescaled both lesion co-occurrence and APM edge weights to the range 0 to 1.

We proceeded to evaluate the community structure of this network with a non-parametric Bayesian hierarchical weighted stochastic block model incorporating layered and attributed properties[96–100] implemented in graph-tool (https://graph-tool. skewed.de).[92–99] We began by fitting a null model, with the two kinds of edge weight—adjusted APM and lesion co-occurrence— randomly distributed across two layers. We then fitted a test model with each type of weight consistently assigned to its own layer. Adjusted APM weights were modelled as Gaussian; lesion co-occurrence weights as Poisson distributions. Having initialised a fit, we used simulated annealing to further optimise it, with a default inverse temperature of 1 to 10.[93] We did not specify a finite number of draws, rather we specified a wait step of 100 iterations for a record-breaking event, to ensure that equilibration was driven by changes in the entropy criterion, instead of driven by a finite number of iterations[99]

We used model entropy to determine if the layered model fit was better than the null, indicating that the inferred community structure distinguished APM and lesion co-occurrence effects.

To visualize the inferred communities, we back projected the incident edge weights onto the brain, deriving the mean and 95% credible intervals for comparison. To examine if modelling lesion co-occurrence requires explicit consideration of aetiology, we replicated the model with the addition of aetiology as a third layer, again conducting formal comparison against a randomized null (Supplementary material).

Synthetic ground truth evaluation

The substrate of a function is definitionally unknown: it is what we are seeking to infer. To examine the comparative fidelity of a set of models we therefore need synthetic ground truths[71] of the complexity likely to obtain in reality. Here we used the meta-analytic repository NeuroQuery[101] to create six realistically complex and distributed ground truth maps across the domains of action, aversion, language, mood, motor and sensation (Supplementary material). The intersection between each lesion and each ground truth was then used to generate a hypothetical deficit for each patient and each domain, and the stochastic block model was subsequently applied exactly as in the case of the real data. The fidelity of the inferred maps was then quantified by their Dice score, and compared to a standard mass-univariate voxel-based lesion-deficit mapping baseline (Supplementary material).

Data and code availability

The data and code that support the findings of this study are available from the corresponding author, L.C., upon reasonable request.

Results

Demographic and behavioural investigations

Frontal, non-frontal and control patients were well-matched for age, gender, chronicity and years of education (all $P > 0.05$). There was no significant difference in lesion volume between LF and RF or between LNF and RNF. There was a significantly greater proportion of tumour patients in frontal than non-frontal [χ^2 (1, $n = 227$) = 5.68, $P < 0.05$] and a significantly greater proportion of stroke patients in non-frontal than frontal [χ^2 (1, $n = 227$) = 7.05, $P < 0.01$]. However, there were no significant differences in the proportion of tumour or stroke patients between LF and RF (all $P > 0.05$) or between LNF and RNF (all $P > 0.05$).

There were no significant differences between frontal, non-frontal and control participants for NART, IL or GNT scores (all $P > 0.05$; Table 1). One-way ANOVAs found highly significant differences between frontal, non-frontal and control participants for S fluency and Hayling suppression errors [$F(2,238) = 25.319$; $P < 0.001$; $F(2,192) = 21.266$; $P < 0.001$, respectively]. Post hoc tests showed that frontal performed significantly worse than non-frontal and control on S fluency ($P < 0.001$; $P < 0.001$, respectively) and Hayling suppression errors ($P < 0.001$; $P < 0.001$, respectively). Pairwise comparisons revealed that LF were significantly more impaired than RF on S fluency ($P < 0.01$), RF were significantly more impaired than LF on Hayling suppression errors ($P < 0.05$; Table 1).

Overall performance

Standard analysis

A one-way ANCOVA controlling for age and NART, found a highly significant difference between frontal, non-frontal and control participants in overall performance [$F(2,387) = 18.491$; $P < 0.001$]. Post hoc tests showed that frontal performed significantly worse than non-frontal ($P < 0.01$) and Control ($P < 0.001$). There was no significant difference between non-frontal and control (corrected $P = 0.185$; Table 2)

Lateralization analysis

A one-way ANCOVA controlling for age and NART found a highly significant difference between LF, RF, LNF, RNF and Control participants [$F(4,385) = 12.237$; $P < 0.001$]. Pairwise comparisons showed a significant difference between RF and Controls ($P < 0.001$) and LF and Control ($P < 0.01$). Importantly, RF were significantly more impaired than LF ($P < 0.01$). There was no significant difference between RNF and Control or LNF and Control; Table 2). Notably, performance fell <1.5 standard deviations (SD) below Control in 43% of RF but only in 22% of LF.

Item difficulty

A 3×3 ANCOVA, controlling for age and NART, revealed a significant main effect of difficulty [$F(2,364) = 5.360$; $P = 0.005$] and Group [$F(2,182) = 22.707$; $P < .001$]. Critically, there was also a significant interaction between difficulty and group [$F(2,364) = 10.822$; $P < 0.001$]. Post hoc pairwise comparisons showed significant differences for the medium group between RF and Control ($P < 0.001$), LF and Control ($P < 0.05$), and RF and LF ($P < 0.05$). Thus, frontal patients were worse than Control, and RF performed the poorest. For the hardest group there were significant differences only between RF and Control ($P < 0.001$), and LF and Control ($P < 0.001$). There were no significant differences for the easy group. Thus, RF impairment on the APM appears to be driven by poor performance on the medium items. Closer inspection revealed that three specific items (3, 5 and 6) were responsible for driving the RF poorer performance than LF, with an accuracy decrement of more than 20% in RF.

Multiple-demand network

A one-way ANCOVA, controlling for age and NART, showed no significant difference in overall performance between patients with

Table 2 Overall performance on APM

	Healthy control	Frontals	Non-frontals	Frontals		Non-frontals	
				Left $n = 69$	Right $n = 77$	Left $n = 39$	Right $n = 42$
Mean number correct/12 (SD)	8.67 (2.41)	7.07[a,**;b,***] (2.78)	8.07 (2.09)	7.71[b,**] (2.49)	6.49[b,***;c,**] (2.93)	8.00 (1.92)	8.14 (2.26)

SD = standard deviation. Scores with significant P-values are in bold.
**$P < 0.01$.
***$P < 0.001$.
[a]Significant difference from non-frontals.
[b]Significant difference from healthy controls.
[c]Significant difference between left and right frontals.

versus without MD lesions [$F(1,172) = 1.88$, $P = 0.172$]. A linear regression analysis, with age, NART and extent of MD involvement entered as predictor variables, significantly predicted APM performance [$r^2 = 0.28$, $F(3,172) = 21.726$, $P < 0.001$]. However, only age and NART (both $P < 0.001$) were significant predictors. Extent of MD involvement did not significantly contribute ($P = 0.410$).

Parcel-based analysis and Bayesian multivariate analysis of multiple-demand network

Parcel-based lesion symptom mapping (PLSM) analyses with age, NART and lesion volume entered as nuisance regressors, revealed that poorer overall performance was associated with right posterior middle frontal gyrus, pars opercularis, precentral gyrus, superior corona radiata and external capsule lesions. When the proportion of MD involvement was entered as a nuisance regressor instead of lesion volume the results remained unchanged.

PLSM analyses on Hayling suppression errors, with age, NART and lesion volume entered as nuisance regressors revealed that poorer performance was associated with right posterior middle frontal gyrus and pars opercularis lesions. When the proportion of MD involvement was entered as a nuisance regressor instead of lesion volume the results remained unchanged.

Bayesian multivariate modelling of individual MD components yielded as credibly predictive only NART (posterior mean coefficient 0.484, 95% credibility interval 0.336 to 0.630), age [mean −0.321, 95% confidence interval (CI) −0.455–−0.187], and a right-sided MD component falling within precentral and posterior medial frontal gyrus (mean −0.310, 95% CI −0.587 to −0.021). The credibility intervals of the coefficients of other MD components all crossed zero.

Network lesion-deficit modelling

Network-based statistics identified a distinct predominantly right frontal network associated with reduced APM (FWER-corrected $P < 0.0001$, t-thresh >3.1) (Fig. 2). The regions with the greatest number of significant nodes (in order of descending degree count) included the right superior frontal gyrus (degree count 22), right middle frontal gyrus (17), right frontal pole (16), right anterior cingulate cortex (16), left superior frontal gyrus (6), right inferior frontal gyrus (4), left anterior cingulate cortex (4), right caudate nucleus (3), right mid cingulate cortex (3), right precentral gyrus (2), right juxtapositional lobule (1), right frontal operculum (1) and right anterior insula (1). A stochastic block model partitioned the network into a structure with three clustered components, broadly encompassing medial wall, superolateral cortical surface and a superior frontal gyrus-dominant component. The addition of lesion aetiology to the list of covariates in the NBS model yielded a near-identical result (test-statistic correlation between significant edges, r.99, $P < 0.0001$) (Supplementary Fig. 2).

Bayesian univariate regression analysis confirmed that patients with lesions overlapping with the network exhibited significantly lower adjusted APM scores, (R^2 0.105, coefficient mean ± SD −0.265 ± 0.068) (95% CI −0.396 to −0.129) (Fig. 3). The extent of overlap, indexed by the degree (i.e. number of nodes) shared between an individual lesion network and the inferred network exhibited a strong log-linear relationship to adjusted APM [R^2 0.190, coefficient mean ± SD −0.002 ± 0.0005 (95% CI −0.00250 to −0.00041]. Bayesian multivariate regression models of adjusted APM predicted by the lesion network adjacency matrix yielded a fit with R^2 0.640. It is important to note that these regression analyses are not independent of the

NBS model: they do not provide further evidence but rather qualify its fidelity.

Generative hierarchical stochastic block modelling of APM performance

The foregoing models inevitably conflate the distributed spatial structure of the underlying neural dependence with that of the causal pathology. To disentangle the two, we need a network model capable of separating the target effects of APM performance from the incidental effects of lesion co-occurrence. This can be achieved with a layered nested stochastic block model, where adjusted APM and lesion co-occurrence weights are distributed in two distinct layers, yielding layer-specific patterns of community structure reflecting the distinct effect of each weight on the network. This model achieved substantially lower entropy—881 118.22 versus 1 182 697.66 nats—than a null model with weights randomised across the two layers (Fig. 4), providing inferential support for distinguishing adjusted APM from co-occurrence effects. This translates to a posterior odds ratio of the layered formulation being $\times 10^{301579}$ more likely than the non-layered null.

The community structure was composed of blocks dominated by adjusted APM, lesion co-occurrence, or neither weight. The adjusted APM layer revealed a set of brain communities with high edge incidence linking the right middle and inferior frontal gyrus (including pars triangularis), right pre- and post-central gyri, and —weakly—the right superior parietal lobule. These communities were sharply distinct from the lesion weight (Figs 4 & 5).

Many of the spatial constraints on the configuration of lesion patterns are imposed by the basic anatomy of the brain and will be shared across aetiologies; those that are not will arise as additional heterogeneity the stochastic block model could theoretically absorb. To determine if aetiology has a substantial structuring effect that merits explicit accounting, we reran the model with an additional, third layer identifying the aetiology of each lesion. This model exhibited far greater description length (2 133 947.48), 2.4× that of the above for an increase of this single feature, indicating a poorer fit to the data and providing no grounds for preference over the simpler model. The anatomical pattern of APM-sensitive communities was in any event very similar (Supplementary Fig. 3).

Synthetic ground truth evaluation

Bayesian model comparison showed all layered models to be more plausible than the null (Supplementary Fig. 4). Compared with VLSM, the stochastic block model achieved significantly superior results across all domains ($P = 0.028$) (Supplementary material and Supplementary Fig. 5). These experiments also demonstrated qualitatively the tendency of VLSM to mislocalize in response to the underlying lesion structure, and the ability of the stochastic block model to resist it.

Discussion

Our study represents the first large-scale investigation of the distributed neural substrates of fluid intelligence in the focally injured brain. We investigated one of the most widely used fluid intelligence tests, the APM, in the largest number of patients with single, focal, unilateral, right or left, frontal or non-frontal lesions and controls. We analysed overall performance, item difficulty and the contribution of MD involvement. For the first time, non-parametric Bayesian stochastic block models were used to reveal the intricate
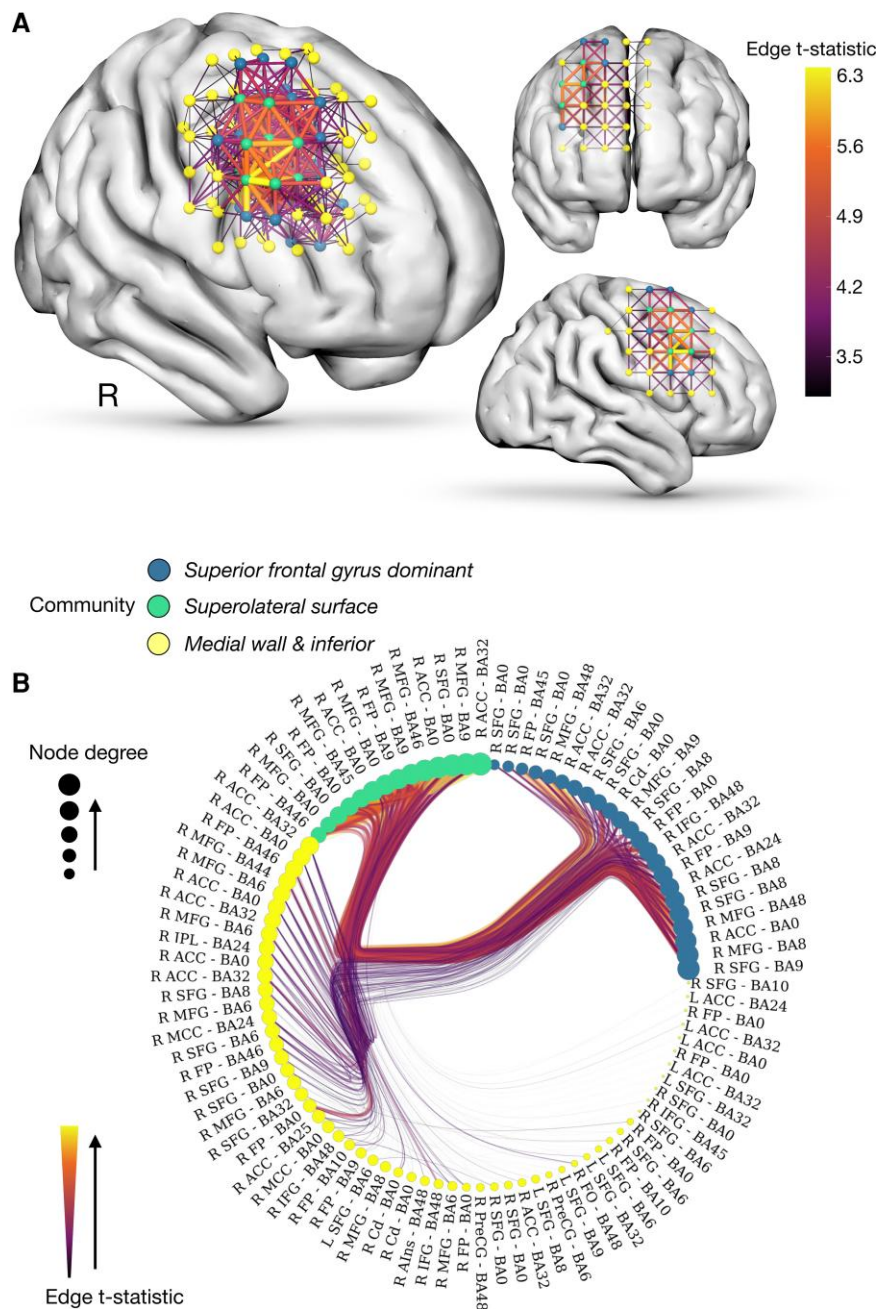
Figure 2 Network modelling of fluid intelligence. (**A**) Network-based statistics identify a significant network associated reduced adjusted APM scores (FWER-P < 0.0001). (**B**) Radial graph of the community structure of the network inferred from a stochastic block model of its statistics shows that the network clusters into three discrete components encompassing the superolateral cortical surface, the medial (and inferior) wall and a superior frontal gyrus dominant cluster. Nodes are colour-coded in accordance with their stochastic block model cluster. Node size is proportional to node degree count. Edge width and colour is proportional to the t-statistic from the model, with a thicker and more yellow line denoting a stronger link between a given network connection and a reduced adjusted APM score. ACG = anterior cingulate gyrus; L = left; IFG = inferior frontal gyrus; IFG-pt = inferior frontal gyrus pars triangularis; MFG = middle frontal gyrus; OFC = orbitofrontal cortex; PreCg = pre-central gyrus; PoCg = post-central gyrus; R = right; SFG = superior frontal gyrus; SPL = superior parietal lobule.

community graph structure of lesion deficit networks, disentangling functional from confounding pathological distributed effects.

Similar to other groups[21–104] and in-keeping with our previous studies[70] we adopted a mixed aetiology approach. Previous comparison of a large frontal and non-frontal sample with different aetiologies on the APM and other executive tests showed that aetiology was not a strong predictor of frontal or non-frontal deficits.[9–105] Hence, different aetiologies do not result in more severe impairments than others and combining across vascular and tumour pathologies is unlikely to significantly distort neuropsychological performance.[75] Instead, focal lesions may relate more closely to the region of damage rather than aetiology. Moreover, data from multiple aetiologies will tend to attenuate distorting effects arising from pathologically driven characteristic patterns of lesion co-occurrence that are widely recognized to bedevil both network and focal lesion-deficit studies. Indeed, less spatial distortion caused by the structure of the
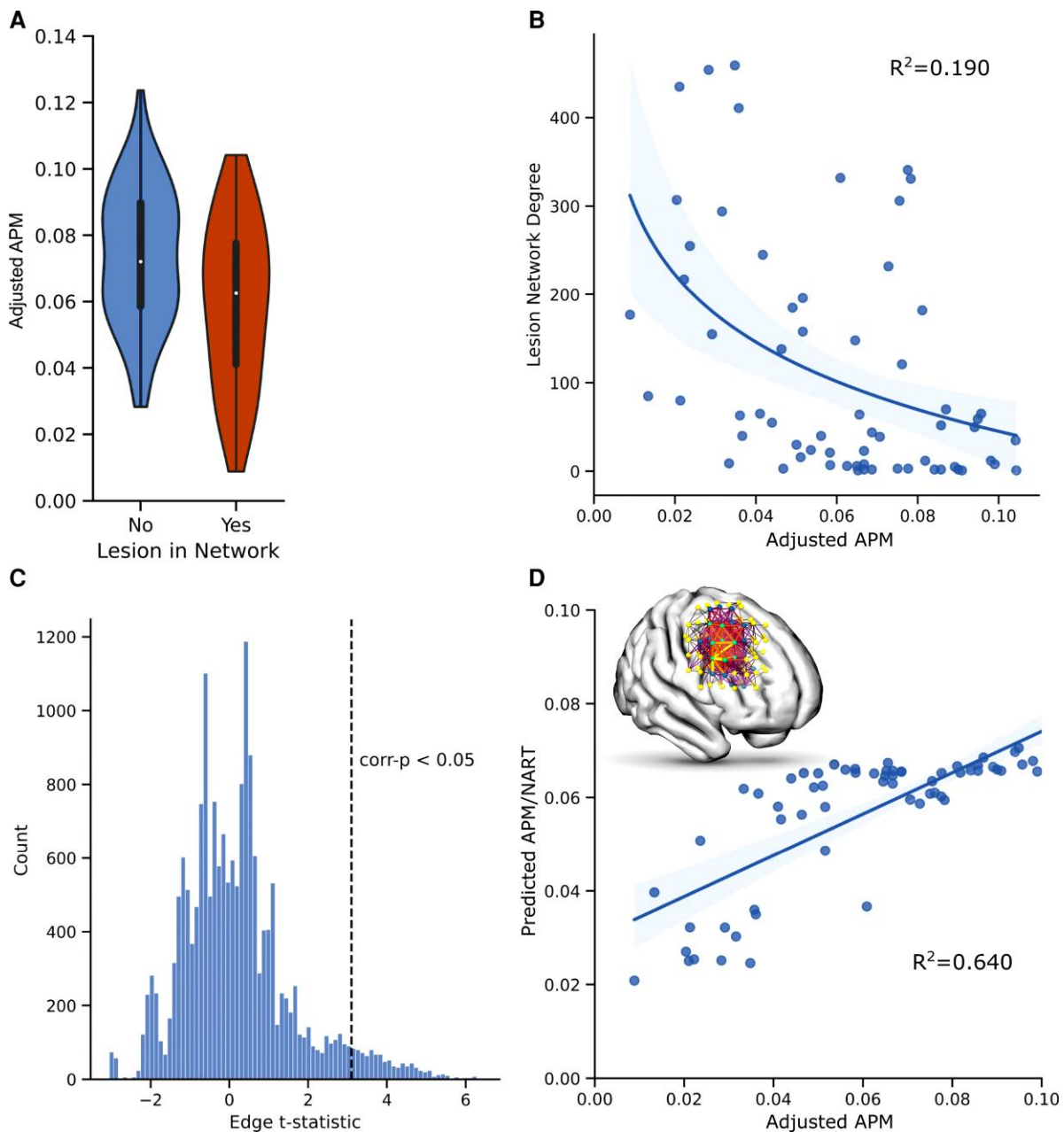
Figure 3 Predicting fluid intelligence from network properties. (A) Violin plots of the adjusted APM scores of patients whose lesions do or do not overlap with the inferred network illustrate significantly lower APM scores in the former [R^2 0.105, coefficient mean $\pm$ SD -0.265 ± 0.068 (95% CI -0.396 to -0.129]. (B) Scatter and line plot shows that the degree count of the overlap of a lesion with the inferred network significantly correlates with adjusted APM scores within a univariate Bayesian regression model [R^2 0.190, coefficient mean $\pm$ SD -0.002 ± 0.0005 (95% CI -0.00250 to -0.00041]. (C) Histogram of the edge *t*-statistics from the network model illustrates the population of edges significantly associated with the APM after multiple comparisons correction. (D) Scatter and line plot shows the predictability of adjusted APM from the network adjacency matrix within a multivariate Bayesian model (R^2 0.640).

pathology may be expected if multiple pathologies differing in their spatial properties are used.

Though fluid intelligence is widely thought to be dependent on the integrity of the frontal lobes, only a handful of focal lesion studies, based on modest samples, have found impairments following frontal lesions.[9,21–55] Applying an array of lesion-deficit models to large scale data, we found APM performance to be specifically vulnerable to the integrity of the right frontal lobe, and largely resistant to damage elsewhere. The left frontal lobe appears to make a contribution to APM performance, if a more modest one. We found

that the performance of the left frontal patients was significantly different from healthy controls and non-frontal patients. However, the left frontal patients performed significantly better than the right frontal patients did.

Our findings speak to the theories of non-frontal involvement in fluid intelligence. The proponents of P-FIT have argued that impairment of fluid intelligence should follow lesions of the posterior and anterior regions that putatively subserve it.[23,24] We found no evidence of such non-frontal causal dependence on APM performance. It is possible that functional imaging findings merely reflect
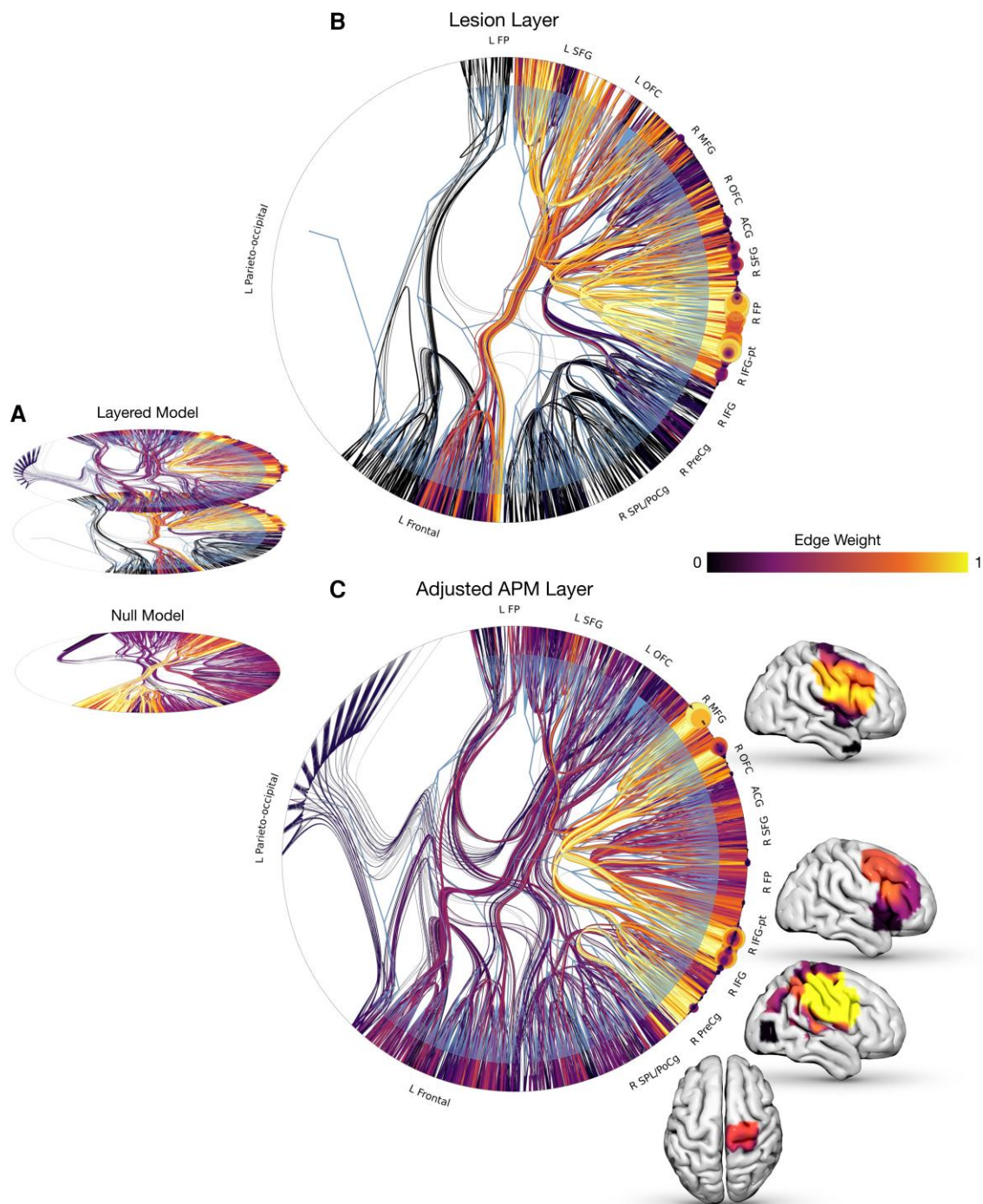
Figure 4 Generative stochastic block models of fluid intelligence. (A) Radial graphs of stochastic block models with adjusted APM and lesion co-occurrence layered (*top*), versus randomly distributed across layers (*bottom*). Edge colour and width is proportional to the associated edge weight. Model entropy favoured the layered over the null model. **(B)** Radial graph illustrating the layered stochastic block model fit with edge colour and width proportional to the lesion co-occurrence weight, and node colour and size proportional to the lesion-weight degree. This demonstrates a community of highly interconnected voxels involving the bilateral frontal pole and orbitofrontal cortex, right superior and inferior frontal gyrus and anterior cingulate gyrus. **(C)** Radial graph illustrating the layered stochastic block model fit with edge colour and width proportional to the adjusted APM weight, node colour and size proportional to the APM-weight degree. This illustrates a characteristically different segregation of brain communities, with high edge incidence linking the right middle and inferior frontal gyrus, (including pars triangularis), right pre-central gyrus and right superior parietal lobule. Brain images are overlayed corresponding to the posterior mean edge weight at these communities. ACG = anterior cingulate gyrus; L = left; IFG = inferior frontal gyrus; IFG-pt = inferior frontal gyrus pars triangularis; MFG = middle frontal gyrus; OFC = orbitofrontal cortex; PreCg = pre-central gyrus; PoCg = post-central gyrus; R = right; SFG = superior frontal gyrus; SPL = superior parietal lobule.
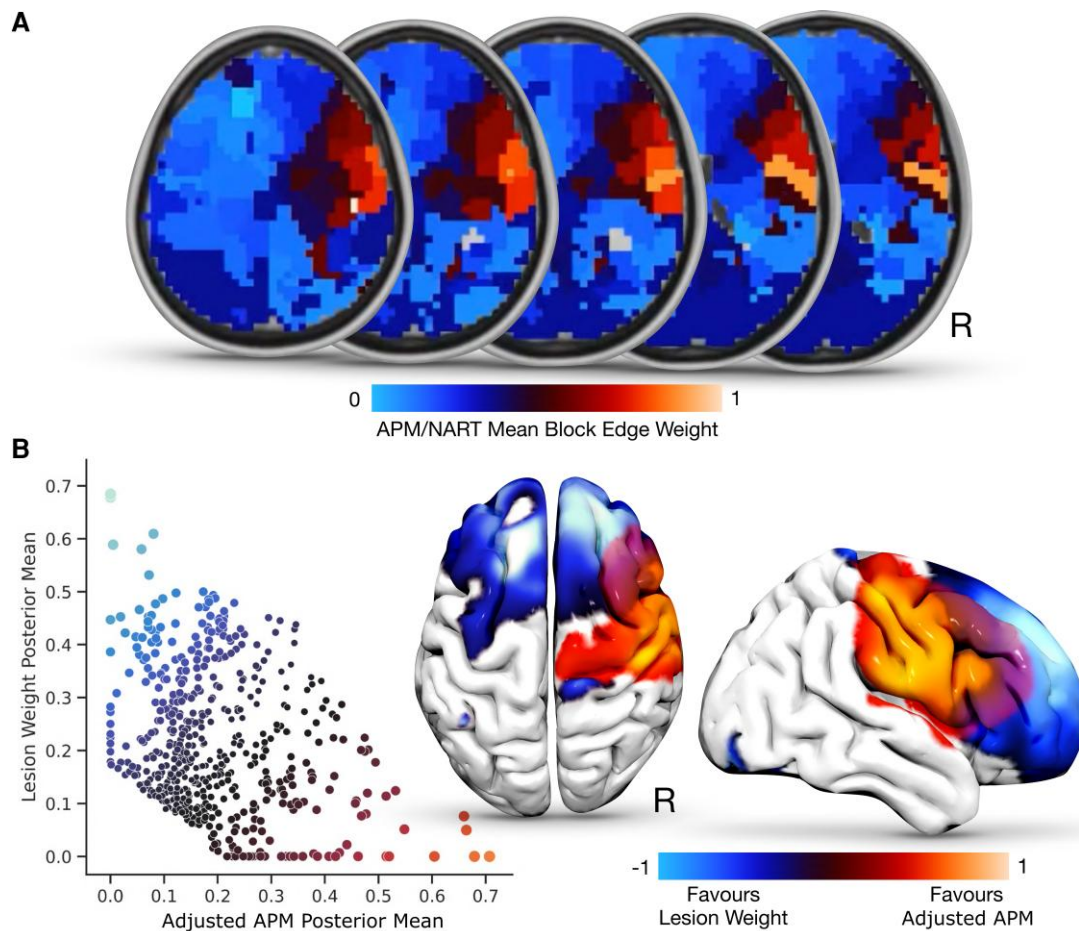
Figure 5 Network communities sensitive to fluid intelligence. A. Axial slices of the mean posterior edge weight for each block at the l_1 aggregation, with more red-orange areas corresponding to a greater value and greater relation to adjusted APM. **B.** Scatterplot illustrating the relationship between posterior mean edge weights at each community block, for both the lesion weight (*y*-axis) and adjusted APM (*x*-axis), with brain reconstructions overlaying these findings. Of note, bilateral frontal-based blocks depicted higher lesion-weight edges, with right fronto blocks more implicating APM. ACG = anterior cingulate gyrus; L = left; IFG= inferior frontal gyrus; IFG-pt = inferior frontal gyrus pars triangularis; MFG = middle frontal gyrus; OFC = orbitofrontal cortex; PreCg = pre-central gyrus; PoCg = post-central gyrus; R = right; SFG = superior frontal gyrus; SPL = superior parietal lobule.

a correlation between fluid intelligence and posterior areas noncritically engaged by the necessary perceptual input.

Our results are also relevant for the MD proposal. Notably, Woolgar *et al.*[79] investigated 80 patients with cortical lesions with a fluid intelligence task (Cattell Culture Fair IQ test). Though the authors reported a significant correlation between MD involvement and fluid intelligence performance overall, in the 44 patients with purely frontal damage the relationship was not significant when non-MD lesion volume was taken into account. So, as far as the frontal lobes are concerned, the authors' theoretical claim was not strongly empirically supported. In our study patients with or without MD damage did not differ significantly in performance on the APM. Moreover, the extent of MD involvement did not contribute to performance. These findings do not support the claim that MD is the seat of fluid intelligence, but neither do they exclude it: absence of evidence is not evidence of absence. In this context, we note that the Woolgar *et al.*[79] study included 30 non-frontal patients. Of these only seven were patients with unilateral parietal lesions, whilst two patients had biparietal lesion. In contrast, our non-frontal sample includes 81 patients, of which 20 have a unilateral parietal lesion. Hence, our study has far greater coverage of non-frontal MD areas than Woolgar's study. While we cannot rule

out the possibility that lower power may be a factor relevant for our conclusion regarding weak contributions from non-frontal MD lesions, our sample is larger than that of the study, which produced the opposite conclusions. Moreover, the Woolgar *et al.* study reported for their parietal patients that MD lesion volume was a significant predictor for performance on fluid intelligence with and without non-MD lesion volume partialled out ($r = -0.65$, $P = 0.042$; $r = -0.63$, $P = 0.035$, respectively). We were not able to replicate this effect in our larger group ($r = -0.063$, $P = 0.811$; $r = -0.17$; $P = 0.950$).

Our findings of greater involvement of the right frontal lobe in APM performance were complemented and extended by our neuroimaging analyses. Both conventional network statistics and nonparametric Bayesian stochastic block modelling heavily implicated the right frontal lobe. Crucially, this localisation was confirmed on explicitly disentangling—uniquely in the field of lesion-deficit mapping—functional from pathology-driven effects within a layered stochastic block model, prominently highlighting a right frontal network including the middle and the inferior frontal gyrus, including pars triangularis, and pre- and post-central gyri, with a comparatively weak contribution from superior parietal lobule. The marked structuring effects of lesion co-occurrence observed

highlight the importance of explicitly modelling them in lesion-deficit inference, whether in the context of network or focal analysis.

Standard PLSM analyses, potentially confounded by lesion co-occurrence effects, suggested that poorer performance was associated with damage to a right frontal network including posterior middle frontal gyrus, pars opercularis, precentral gyrus, superior corona radiata and external capsule, invariantly to the degree of MD involvement. That a similar set of RF regions were implicated in Hayling suppression errors, a verbal test, suggests that function lateralization in the frontal lobes is not explained by task sensitivity to language alone.

Behaviourally we found a highly significant interaction between item difficulty and frontal lesion lateralisation. The asymmetry in performance was nearly three times greater for the middle than for the highest level of difficulty, with neither population nearing the ceiling or floor. Why might that be?

While complete agreement is lacking, factor analyses of progressive matrices indicate at least two material components. Dillon et al.[106] identify a factor related to 'perceiving the progression of a pattern' (p.1301), and another to 'the addition and/or subtraction of elements'. Lynn et al.[107] offer 'Gestalt continuation', following Van der Ven and Ellis,[108] and 'verbal-analytic reasoning', respectively. It is apparent that the three medium items (3, 5, 6) of the APM showing the greatest lateralisation, are all those where perceiving the progression of a pattern is an obvious approach. By contrast, the non-ceiling items with the smallest lateralization (10, 11, 12) are all those where addition or subtraction come into play. Though the limited number of items precludes firm conclusions on factorisation, these findings suggest the components of APM may lateralise in different ways. Inferences involving the perception of a progressive pattern may be especially sensitive to the integrity of the right frontal network.

In conclusion, our study represents the most robust investigation of the hitherto poorly characterized fluid intelligence in patients with single, focal, unilateral lesions. Our approach of combining novel graph-based lesion-deficit mapping with detailed investigation of APM performance in a large sample of patients provides crucial information about the neural basis of fluid intelligence. We suggest that a right frontal network, rather than a wide set of regions distributed across the brain, is critical to the high-level inferences, based on perceiving pattern progression, involved in fluid intelligence. Our findings further corroborate the clinical utility of APM in evaluating fluid intelligenceand identifying right frontal lobe dysfunction.

Funding

Competing interests

The authors report no competing interests.

Supplementary material

Supplementary material is available at *Brain* online.

References

1. Varriale V, van der Molen MW, De Pascalis V. Mental rotation and fluid intelligence: a brain potential analysis. *Intelligence*. 2018;69:146-157.
2. Chan E, MacPherson SE, Bozzali M, Shallice T, Cipolotti L. The influence of fluid intelligence, executive functions and pre-morbid intelligence on memory in frontal patients. *Front Psychol*. 2018;9:926.
3. Plomin R, Deary IJ. Genetics and intelligence differences: five special findings. *Mol Psychiatry*. 2015;20:98-108.
4. Deary IJ, Taylor MD, Hart CL, et al. Intergenerational social mobility and mid-life status attainment: influences of childhood intelligence, childhood social factors, and education. *Intelligence*. 2005;33:455-472.
5. Gottfredson LS, Deary IJ. Intelligence predicts health and longevity, but why? *Curr Dir Psychol Sci*. 2004;13:1-4.
6. Deary I. Why do intelligent people live longer? *Nature*. 2008;456:175-176.
7. Shallice T, Cipolotti L. The prefrontal cortex and neurological impairments of active thought. *Annu Rev Psychol*. 2018;69:157-180.
8. Whalley LJ, Deary IJ, Appleton CL, Starr JM. Cognitive reserve and the neurobiology of cognitive aging. *Ageing Res Rev*. 2004;3:369-382.
9. Cipolotti L, Healy C, Chan E, et al. The impact of different aetiologies on the cognitive performance of frontal patients. *Neuropsychologia*. 2015;68:21-30.
10. Kievit RA, van Rooijen H, Wicherts JM, et al. Intelligence and the brain: a model-based approach. *Cognit Neurosci*. 2012;3:89-97.
11. Carroll JB. *Human cognitive abilities: a survey of factor-analytic studies*. Cambridge University Press; 1993.
12. Personality If, Testing A. *Measuring intelligence with the culture fair tests*. The Institute for Personality and Ability Testing; 1973.
13. Raven JC. *Manual for the advanced progressive matrices*. Set, 1; 1976.
14. Carpenter PA, Just MA, Shell P. What one intelligence test measures: a theoretical account of the processing in the raven progressive matrices test. *Psychol Rev*. 1990;97:404-431.
15. Heim A. *Manual for the AH4 group test of general intelligence*. National Foundation for Education Research; 1970.
16. Wechsler D. *WAIS-III administration and scoring manual*. The Psychological Corporation; 1997.
17. Cattell RB, Horn JL. A check on the theory of fluid and crystallized intelligence with description of new subtest designs. *J Educ Meas*. 1978;15:139-164.
18. Sternberg RJ. *In search of the human mind*. Harcourt Brace College; 1995.
19. Marshalek B, Lohman DF, Snow RE. The complexity continuum in the radex and hierarchical models of intelligence. *Intelligence*. 1983;7:107-127.
20. Duncan J, Burgess P, Emslie H. Fluid intelligence after frontal lobe lesions. *Neuropsychologia*. 1995;33:261-268.
21. Roca M, Parr A, Thompson R, et al. Executive function and fluid intelligence after frontal lobe lesions. *Brain*. 2010;113:234-247.
22. Duncan J, Seitz RJ, Kolodny J, et al. A neural basis for general intelligence. *Science*. 2000;289:457-460.
23. Colom R, Karama S, Jung RE, Haier RJ. Human intelligence and brain networks. *Dialogues Clin Neurosci*. 2010;12:489-501.

24. Jung RE, Haier RJ. The parieto-frontal integration theory (P-FIT) of intelligence: converging neuroimaging evidence. *Behav Brain Sci.* 2007;30:135-154.

25. Dubois J, Galdi P, Paul LK, Adolphs R. A distributed brain network predicts general intelligence from resting-state human neuroimaging data. *Philos Trans R Soc Lon B Biol Sci.* 2018;373:20170284.

26. Prabhakaran V, Rypma B. P-FIT and the neuroscience of intelligence: how well does P fit? *Behav Brain Sci.* 2007;30:166-167.

27. Prabhakaran V, Rypma B, Gabrieli JD. Neural substrates of mathematical reasoning: a functional magnetic resonance imaging study of neocortical activation during performance of the necessary arithmetic operations test. *Neuropsychology.* 2001;15:115.

28. Santarnecchi E, Emmendorfer A, Pascual-Leone A. Dissecting the parieto-frontal correlates of fluid intelligence: a comprehensive ALE meta-analysis study. *Intelligence.* 2017;63:9-28.

29. Corbetta M, Patel G, Shulman GL. The reorienting system of the human brain: from environment to theory of mind. *Neuron.* 2008;58:306-324.

30. Chen Z, De Beuckelaer A, Wang X, Liu J. Distinct neural substrates of visuospatial and verbal-analytic reasoning as assessed by raven's advanced progressive matrices. *Sci Rep.* 2017;7:16230.

31. Prabhakaran V, Smith JAL, Desmond JE, Glover GH, Gabrieli JDE. Neural substrates of fluid reasoning: an fMRI study of neocortical activation during performance of the raven's progressive matrices test. *Cognit Psychol.* 1997;33:43-63.

32. Gilaie-Dotan S, Saygin AP, Lorenzi LJ, Rees G, Behrmann M. Ventral aspect of the visual form pathway is not critical for the perception of biological motion. *Proc Natl Acad Sci.* 2015;112:E361-E370.

33. Croquelois A, Bogousslavsky J. Stroke aphasia: 1,500 consecutive cases. *Cerebrovasc Dis.* 2011;31:392-399.

34. Jung-Beeman M. Bilateral brain processes for comprehending natural language. *Trends Cogn Sci.* 2005;9:512-518.

35. Bakare CGM, Adeloye A. Intellectual deficits in Nigerians after missile wounds of the brain. *Brain.* 1972;95:79-86.

36. Barbey AK, Colom R, Grafman J. Dorsolateral prefrontal contributions to human intelligence. *Neuropsychologia.* 2013;51:1361-1369.

37. Black FW. Cognitive deficits in patients with unilateral war-related frontal lobe lesions. *J Clin Psychol.* 1976;32:366-372.

38. Grafman J, Salazar A, Weingartner H, Vance S, Amin D. The relationship of brain-tissue loss volume and lesion location to cognitive deficit. *J Neurosci.* 1986;6:301-307.

39. Newcombe F. *Missile wounds of the brain: a study of psychological deficits.* Oxford University Press; 1969.

40. Salazar AM, Grafman J, Schlesselman S, *et al.* Penetrating war injuries of the basal forebrain: neurology and cognition. *Neurology.* 1986;36:459-465.

41. Teuber H-L, Battersby WS, Bender MB. Effects of cerebral lesions on intellectual functioning in man. *Fed Proc.* 1952;11:161.

42. Teuber H-L. Some alterations in behavior after cerebral lesions in man. In: Bass AD, eds. *Evolution of nervous control from primitive organisms in man*: American Association for the Advancement of Science. 1959:157–194.

43. Weinstein S, Teuber L. Effects of penetrating brain injury on intelligence test scores. *Science.* 1957:125:1036-1037.

44. Barbey AK, Colom R, Paul EJ, Grafman J. Architecture of fluid intelligence and working memory revealed by lesion mapping. *Brain Struct Funct.* 2014;219:485-494.

45. Miklewska A, Kaczmarek M, Strelau J. The relationship between temperament and intelligence: cross-sectional study in successive age groups. *Pers Individ Differ.* 2006;40:643-654.

46. Jolly AE, Bălăeț M, Azor A, *et al.* Detecting axonal injury in individual patients after traumatic brain injury. *Brain.* 2021;144:192-113.

47. Barbey AK, Colom R, Solomon J, Krueger F, Forbes C, Grafman J. An integrative architecture for general intelligence and executive function revealed by lesion mapping. *Brain.* 2012;135:1154-1164.

48. Gläscher J, Rudrauf D, Colom R, *et al.* Distributed neural system for general intelligence revealed by lesion mapping. *Proc Natl Acad Sci U S A.* 2010;107:4705-4709.

49. Bowren M Jr, Adolphs R, Bruss J, *et al.* Multivariate lesion-behavior mapping of general cognitive ability and its psychometric constituents. *J Neurosci.* 2020;40:8924-8937.

50. Gläscher J, Tranel D, Paul LK, *et al.* Lesion mapping of cognitive abilities linked to intelligence. *Neuron.* 2009;61:681-691.

51. Warrington EK, James M, Maciejewski C. The WAIS as a lateralizing and localizing diagnostic instrument: a study of 656 patients with unilateral cerebral lesions. *Neuropsychologia.* 1986;24:223-239.

52. Tranel D, Manzel K, Anderson SW. Is the prefrontal cortex important for fluid intelligence? A neuropsychological study using matrix reasoning. *Clin Neuropsychol.* 2008;22:242-261.

53. Eslinger PJ, Damasio AR. Severe disturbance of higher cognition after bilateral frontal lobe ablation: patient EVR. *Neurology.* 1985;35:1731-1741.

54. Shallice TIM, Burgess PW. Deficits in strategy application following frontal lobe damage in man. *Brain.* 1991;114:727-741.

55. Mole J, Foley J, Shallice T, Cipolotti L. The left frontal lobe is critical for the AH4 fluid intelligence test. *Intelligence.* 2021;87:101564.

56. Archibald M, Wepman JM, Jones LV. Nonverbal cognitive performance in aphasic and nonaphasic brain-damaged patients. *Cortex.* 1967;3:275-294.

57. Basso A, De Renzi E, Faglioni P, Scotti G, Spinnler H. Neuropsychological evidence for the existence of cerebral areas critical to the performance of intelligence tasks. *Brain.* 1973;96:715-728.

58. Colonna A, Faolioni P. The performance of hemisphere-damaged patients on spatial intelligence tests. *Cortex.* 1966; 2:293-307.

59. Costa LD, Vaughan HG Jr. Performance of patients with lateralized cerebral lesions. I. Verbal and perceptual tests. *J Nerv Ment Dis.* 1962;134:162-168.

60. De Renzi E, Faglioni P. The comparative efficiency of intelligence and vigilance tests in detecting hemispheric cerebral damage. *Cortex.* 1965;1:410-433.

61. Denes F, Semenza C, Stoppa E, Gradenigo G. Selective improvement by unilateral brain-damaged patients on raven coloured progressive matrices. *Neuropsychologia.* 1978;16:749-752.

62. Arrigoni G, De Renzi E. Constructional apraxia and hemispheric locus of lesion. *Cortex.* 1964;1:170-197.

63. Kertesz A, McCabe P. Intelligence and aphasia: performance of aphasics on raven's coloured progressive matrices (RCPM). *Brain Lang.* 1975;2:387-395.

64. Costa LD. Interset variability on the raven coloured progressive matrices as an indicator of specific ability deficit in brain-lesioned patients. *Cortex.* 1976;12:31-40.

65. Costa LD, Vaughan HG Jr, Horwitz M, Ritter W. Patterns of behavioral deficit associated with visual spatial neglect. *Cortex.* 1969;5:242-263.

66. Villardita C. Raven's colored progressive matrices and intellectual impairment in patients with focal brain damage. *Cortex.* 1985; 21:627-634.

67. Basso A, Capitani E, Luzzatti C, Spinnler H. Intelligence and left hemisphere disease. The role of aphasia, apraxia and size of lesion. *Brain.* 1981;104:721-734.

68. Boller F, Vignolo LA. Latent sensory aphasia in hemisphere-damaged patients: an experimental study with the token test. *Brain*. 1966;89:815-830.

69. Zangwill OL. Intelligence in aphasia. In: De Reuck AVS, O'Connor M, eds. *Disorders of language*. Churchill. 1964:261-274.

70. Cipolotti L, Xu T, Harry B, *et al*. Multi-model mapping of phonemic fluency. *Brain Commun*. 2021;3: fcab232.

71. Mah YH, Husain M, Rees G, *et al*. Human brain lesion-deficit inference remapped. *Brain*. 2014;137:2522-2531.

72. Warrington E, James M. *The visual object and space perception battery*. Pearson; 1991.

73. Mckenna P, Warrington E. *Graded naming test*. NFER-Nelson; 1983.

74. Nelson H. *National adult reading test (NART)*. NFER-Nelson; 1982.

75. Cipolotti L, Molenberghs P, Dominguez J, *et al*. Fluency and rule breaking behaviour in the frontal cortex. *Neuropsychologia*. 2020;137:107308.

76. Spreen O, Strauss EA. *A compendium of neuropsychological tests: administration, norms, and commentary*. Oxford University Press; 1998.

77. Burgess PW, Shallice T. *The hayling and brixton tests*. Thames Valley Test Company; 1997.

78. Cipolotti L, MacPherson SE, Gharooni S, *et al*. Cognitive estimation: performance of patients with focal frontal and posterior lesions. *Neuropsychologia*. 2018;115:70-77.

79. Woolgar A, Parr A, Cusack R, *et al*. Fluid intelligence loss linked to restricted regions of damage within frontal and parietal cortex. *Proc Natl Acad Sci U S A*. 2010;107:14899-14902.

80. Faria AV, Joel SE, Zhang Y, *et al*. Atlas-based analysis of resting-state functional connectivity: evaluation for reproducibility and multi-modal anatomy–function correlation studies. *Neuroimage*. 2012;61:613-621.

81. Winkler AM, Ridgway GR, Webster MA, Smith SM, Nichols TE. Permutation inference for the general linear model. *Neuroimage*. 2014;92:381-397.

82. Watanabe S. A widely applicable Bayesian information criterion. *J Mach Learn Res*. 2013;14:867-897.

83. Gordon EM, Laumann TO, Adeyemo B, Huckins JF, Kelley WM, Petersen SE. Generation and evaluation of a cortical area parcellation from resting-state correlations. *Cereb Cortex*. 2016; 26:288-303.

84. Glasser MF, Coalson TS, Robinson EC, *et al*. A multi-modal parcellation of human cerebral cortex. *Nature*. 2016;536:171-178.

85. Tzourio-Mazoyer N, Landeau B, Papathanassiou D, *et al*. Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *Neuroimage*. 2002;15:273-289.

86. Desikan RS, Segonne F, Fischl B, *et al*. An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *Neuroimage*. 2006; 31:968-980.

87. Zalesky A, Fornito A, Bullmore ET. Network-based statistic: identifying differences in brain networks. *Neuroimage*. 2010; 53:1197-1207.

88. Zalesky A, Cocchi L, Fornito A, Murray MM, Bullmore E. Connectivity differences in brain networks. *Neuroimage*. 2012; 60:1055-1062.

89. Bullmore E, Sporns O. Complex brain networks: graph theoretical analysis of structural and functional systems. *Nat Rev Neurosci*. 2009;10:186-198.

90. Freedman D, Lane D. A nonstochastic interpretation of reported significance levels. *J Bus Econ Stat*. 1983;1:292-298.

91. Sporns O, Tononi G, Kötter R. The human connectome: a structural description of the human brain. *PLoS Comput Biol*. 2005;1: e42.

92. Peixoto TP. Nonparametric weighted stochastic block models. *Phys Rev E*. 2018;97:012306.

93. Peixoto TP. Efficient monte carlo and greedy heuristic for the inference of stochastic block models. *Phys Rev E*. 2014;89: 012804.

94. Makalic E, Schmidt DF. High-Dimensional Bayesian regularized regression with the bayesreg package. *arxiv*. [Preprint] doi:10.48550/arXiv.1611.06649

95. Schmidt DF. Bayesreg. https://cran.r-project.org/web/packages/bayesreg/bayesreg.pdf

96. Peixoto TP. Inferring the mesoscale structure of layered, edge-valued, and time-varying networks. *Phys Rev E*. 2015; 92:042807.

97. Vaca-Ramirez F, Peixoto TP. Systematic assessment of the quality of fit of the stochastic block model for empirical networks. *Phys Rev E*. 2022;105:054311.

98. Faskowitz J, Yan X, Zuo X-N, Sporns O. Weighted stochastic block models of the human connectome across the life span. *Sci Rep*. 2018;8:12997.

99. Ruffle JK, Hyare H, Howard MA, *et al*. The autonomic brain: multi-dimensional generative hierarchical modelling of the autonomic connectome. *Cortex*. 2021;143:164-179.

100. Adolphs R. Human lesion studies in the 21st century. *Neuron*. 2016;90:1151-1153.

101. Dockes J, Poldrack RA, Primet R, *et al*. Neuroquery, comprehensive meta-analysis of human brain mapping. *Elife*. 2020;9: e53385.

102. Gläscher J, Adolphs R, Damasio H, *et al*. Lesion mapping of cognitive control and value-based decision making in the prefrontal cortex. *Proc Natl Acad Sci U S A*. 2012;109: 14681-14686.

103. Inoue K, Madhyastha T, Rudrauf D, Mehta S, Grabowski T. What affects detectability of lesion–deficit relationships in lesion studies? *Neuroimage Clin*. 2014;6:388-397.

104. Urbanski M, Bréchemier ML, Garcin B, *et al*. Reasoning by analogy requires the left frontal pole: lesion-deficit mapping and clinical implications. *Brain*. 2016;139:1783-1799.

105. MacPherson SE, Allerhand M, Gharooni S, *et al*. Cognitive reserve proxies do not differentially account for cognitive performance in patients with focal frontal and non-frontal lesions. *J Int Neuropsychol Soc*. 2020;26:739-748.

106. Dillon RF, Pohlmann JT, Lohman DF. A factor analysis of raven's advanced progressive matrices freed of difficulty factors. *Educ Psychol Meas*. 1981;41:1295-1302.

107. Lynn RJ, Allik J, Irwing P. Sex differences on three factors identified in raven's standard progressive matrices. *Intelligence*. 2004;32:411-424.

108. Van der Ven AHGS, Ellis JL. A rasch analysis of raven's standard progressive matrices. *Pers Individ Differ*. 2000;29:45-64.