anales de psicología / annals of psychology
2026, vol. 42, nº 2 (may-august), e23, 1-12
https://doi.org/10.6018/analesps.673191

Published by Servicio de Publicaciones, Universidad de Murcia (Spain),
in https://revistas.um.es/analesps

Quote the Raven? Always! A brain morphometric analysis of the intelligence hierarchy

Francisco J. Román[1,*], Sergio Escorial[2], Miguel Burgaleta[3], Kenia Martínez[1], and Roberto Colom[1]

1 Facultad de Psicología, Universidad Autónoma de Madrid, Madrid, Spain
2 Facultad de Psicología, Universidad Complutense de Madrid, Madrid, Spain
3 Facultat de Psicologia, Universitat de Barcelona, Barcelona, Spain

Título: ¿Citar al Raven? ¡Siempre! Un análisis morfométrico cerebral de la jerarquía de la inteligencia

Resumen: En este estudio se comprueba la hipótesis de que, si el test de matrices progresivas de Raven (RAPM) es la mejor medida disponible del factor general de inteligencia (*g*), entonces sus correlatos cerebrales deben mostrar una mayor coincidencia con *g* que la observada en otras medidas de inteligencia. Para ello, se analizaron los correlatos de materia gris con los diferentes niveles de la jerarquía de la inteligencia, comprobando las consistencias y discrepancias entre los test y los constructos. Un grupo de 104 adultos jóvenes completó una batería de nueve pruebas de inteligencia que medían la inteligencia fluida-abstracta (*Gf*), cristalizada-verbal (*Gc*) y espacial (*Gv*). Este grupo también se realizó una resonancia magnética. Se calcularon las puntuaciones psicométricas para el factor de orden superior que representaba la inteligencia general (*g*), los factores de grupo de primer orden que representaban *Gf, Gc y Gv* y, por último, las nueve medidas específicas. Se obtuvieron evaluaciones optimizadas de morfometría basada en vóxeles (VBM) y morfometría basada en la superficie cortical (SBM, grosor cortical y área de la superficie cortical) y sus indicadores se relacionaron con las puntuaciones. Se inspeccionaron sistemáticamente las superposiciones entre los mapas cerebrales resultantes, controlando el sexo, la edad y la lateralidad. Los resultados de VBM identificaron una superposición en un grupo del giro frontal medio derecho, mientras que los resultados del área de la superficie cortical mostraron una superposición en la corteza prefrontal dorsolateral derecha, para RAPM, *Gf* y *g*. Estos resultados se consideraron coherentes con la hipótesis principal y, por lo tanto, respaldan al RAPM como la mejor estimación única de *g*.

Palabras clave: Inteligencia. Morfometría basada en vóxeles (VBM). Morfometría basada en la superficie (SBM). Test de matrices progresivas avanzadas de Raven (RAPM). Lóbulo frontal derecho.

Abstract: Here we test the hypothesis that if the Raven Progressive Matrices Test (RAPM) is the best available single measure of the general factor of intelligence (*g*), then their brain correlates must show a greater overlap with *g* than that observed for other intelligence measures. For that purpose, gray matter correlates were analyzed at different levels of the intelligence hierarchy checking for consistencies and discrepancies among measurements and constructs. A group of 104 young adults completed a battery of nine intelligence tests measuring fluid-abstract (*Gf*), crystallized-verbal (*Gc*), and spatial (*Gv*) intelligence. They also undertook MRI scanning. Psychometric scores were computed for a higher-order factor representing general intelligence (*g*), first-order group factors representing *Gf, Gc*, and *Gv*, and, finally, the nine specific measures. Optimized voxel-based morphometry (VBM) and cortical surface-based morphometry (SBM, cortical thickness and cortical surface area) assessments were obtained and related to the scores. Overlaps among resulting brain maps were systematically inspected, controlling for sex, age, and handedness. VBM findings identified an overlap in a right middle frontal gyrus cluster, whereas cortical surface area results showed a key overlap in the right dorsolateral prefrontal cortex, for the RAPM, *Gf*, and *g*. These results were seen as consistent with the main hypothesis and, therefore, support the RAPM as the best single estimate of *g*.

Keywords: Intelligence. Voxel-based morphometry (VBM). Surface-based morphometry (SBM). Raven Advanced Progressive Matrices Test (RAPM). Right frontal lobe.

Introduction

The Raven's Advanced Progressive Matrices test (RAPM, Raven, 1939) is administered worldwide for measuring general intelligence or *g*. Available evidence supports its psychometric properties (Alderton & Larson, 1990; Arthur & Woehr, 1993; Dillon et al., 1981; Dolke, 1976; Lynn et al., 2004; Torres & Cuesta, 1992; Winfred & Day, 1994), its unbiased status for measuring population differences on a global level (Abad et al., 2004; Colom & Abad, 2005; Colom et al., 2004; Colom & García-López, 2002; Mackintosh & Bennett, 2005) and its adequacy for cross-cultural studies (e.g. Wicherts et al., 2010). This standardized intelligence measure is used both in research and applied settings (e.g., Muñiz & Fernández-Hermida, 2010; Raven & Raven, 2008) and has recently been used to test how "intelligent" are Artificial Intelligence systems (Małkiński & Mańdziuk, 2025; Mitchell, 2021; Webb et al., 2023; Zhao et al., 2023).

The RAPM was originally designed for measuring inductive reasoning, which is a key facet of fluid intelligence (*Gf*) (Carroll, 1993; Haier et al., 2023). It has been accepted as one the best available measurements of the general factor of intelligence or *g* (Deary & Stough, 1996; Deshon et al., 1995; Jensen, 1998; McGrew & Flanagan, 1998; Spearman, 1938; Spearman & Wynn-Jones, 1951). In his encyclopedic book, Jensen (1998) concluded: "when the Progressive Matrices test is factor analyzed along with a variety of other tests it is typically among the two or three tests having the highest *g* loading, usually around .80. Probably its most distinct feature is its very low loadings on any factor other than *g*" (p. 38).

However, these statements did not carry a general consensus. For example, in his seminal paper *"Quote the Raven? Nevermore!"* Hunt (1974) argued that several items of the Raven test can be approached by perceptual processes which do not involve the analytical reasoning required for tapping the core of the intelligence construct (Spearman & Wynn-Jones, 1951). It has been argued that the Progressive Matrices Test cannot be regarded as the best marker of general intelligence because of its significant visuospatial component (Ackerman et al., 2002; Burke, 1958; see Gonthier, 2022). Indeed, it has been maintained that its predictive validity is

---

*** Correspondence address [Dirección para correspondencia]:**
Francisco J. Román. Facultad de Psicología, Universidad Autónoma de Madrid, 28049, Madrid (Spain) Email: f.javier.roman@uam.es
(Article received: 24-07-25, revised: 07-11-2025; accepted: 13-11-2025)

smaller than the predictive validity achieved via batteries such as the Weschler scales or the Stanford-Binet (Ackerman, 1999; Ackerman et al., 2002). Jensen (1980) himself acknowledged that the Raven test is not the best indicator of general intelligence (or *Gf*) in a response to Cattell (1980) accepting that the "Cattell's Culture Fair Intelligence Test" is better for measuring fluid intelligence because it includes a variety of cognitive challenges beyond the matrix format.

Generally speaking, psychometric and cognitive research raises some doubts regarding the special status of the Raven Progressive Matrices Test for measuring key facets of the *Gf-g* construct. Nevertheless, the Raven is often used to assess the general factor of intelligence in related research fields such as neuroscience (Colom et al., 2010; Colom & Thompson, 2011; Gray et al., 2003; Haier et al., 2023). In neuroimaging studies, the Standard Progressive Matrices (SPM) or the more difficult Raven Advanced Progressive Matrices (RAPM) have been used to identify brain correlates of intelligence (Gray et al., 2003; Haier et al., 1988, 2003; Lee et al., 2006; O'Boyle et al., 2005).

A classic PET study by Prabhakaran et al (1997) found distinguishable activity patterns in the brain for the different rules underlying RAPM items, as defined by Carpenter et al. (1990). For example, the right frontal and bilateral parietal regions were more activated for figural items than for match items (control items); bilateral frontal and left parietal, occipital, and temporal regions were more activated for analytical than for figural items. These researchers concluded that the neural network underlying Raven performance overlaps verbal working memory networks, even when considering analytical reasoning with nonverbal patterns. Therefore, strong links were detected between neural systems underlying working memory and fluid reasoning.

Recent *f*MRI research has used the Raven's Standard Progressive Matrices (RSPM) to obtain temporal and functional insights into the neural basis of intelligence (Zurrin et al., 2024). Using a dimensionality reduction approach, five large-scale task-related BOLD networks were identified. The multiple-demand network (MDN) showed early activation (~9 s) during solution searching, followed by the response selection (RESP) network (~12 s) and the re-evaluation (RE-EV) network (~18 s), associated with solution checking. A positive correlation between MDN and RE-EV activity suggests coordinated processing. Higher MDN activation was linked to lower accuracy on difficult items, indicating inefficient resource use in lower performers. These results support and refine the Parieto-Frontal Integration Theory (P-FIT; Jung & Haier, 2007).

Here we analyze regional gray matter correlates of several intelligence tests measuring fluid, crystallized, and spatial intelligence. As recommended, but rarely properly addressed, each psychological construct comprised three specific measures. Brain structural correlates were systematically analyzed at different levels of the intelligence hierarchy (Colom et al., 2009; Colom & Thompson, 2011; Haier et al., 2009a; Martínez & Colom, 2021): (a) specific measures, (b) first-order group factors, and (c) the higher-order factor (*g*). The main hypothesis is that if the Progressive Matrices Test is indeed the best psychological specific measure of the general factor of intelligence (*g*), then its gray matter correlates will show a unique and specific overlap with first-order *Gf* and higher-order *g*. Further, this specific overlap will tap parieto-frontal regions, as proposed by the Parieto-Frontal integration theory (P-FIT) of intelligence (Jung & Haier, 2007) and would not be found for other specific measures or group factors.

Method

Participants

405 university undergraduates from the Universidad Autónoma de Madrid and the Universidad Complutense de Madrid completed a battery of nine tests measuring intelligence. 120 participants representative of the range of test scores were invited for MRI scanning (50% females). 104 agreed to join in the study (59 females and 45 males, mean age=19.9, SD=1.6, age range=18 to 27; 93.3% right-handed). All participants gave informed written consent. Participants completed a questionnaire asking for medical, neurological, and psychiatric illness, or conditions that might preclude MRI scanning. They received a payment of 20 € for participation.

Intelligence measures

Intelligence was measured by nine tests tapping abstract-fluid (*Gf*), verbal-crystallized (*Gc*), and spatial intelligence (*Gv*) (Horn, 1985). The administered tests were the RAPM (Raven, 1962) (screening version, even numbered items), three subtests from the *Primary Mental Abilities Battery* (PMA; Thurstone & Thurstone, 1968), namely, inductive reasoning (R), vocabulary (V), and mental rotation (S), four subtests from the *Differential Aptitude Test Battery* (DAT-5) (Bennett et al., 1990), specifically screening versions (even numbered items) for the abstract reasoning (AR), verbal reasoning (VR), spatial relations (SR) and numerical reasoning (NR) subtests, and, finally, the *Rotation of solid figures test* (Yela, 1969).

Three tests were used to assess *Gf*. The *Raven's Advanced Progressive Matrices* (RAPM) consists of 36 matrix items, each requiring participants to identify the missing entry in a 3×3 matrix from eight alternatives. Total score is the number of correct responses, with a 20-minute time limit. The *Differential Aptitude Test – Abstract Reasoning* (DAT-AR) includes 20 items based on abstract figures; each item presents a visual rule to be completed by selecting one of five alternatives. Administration time was 10 minutes. The *Primary Mental Abilities – Reasoning* (PMA-R) consists of 30 letter series items (e.g., a–c–a–c–...), where participants must select the correct next letter from six options. The test duration was 6 minutes.

Gc was measured using three tests. The *Differential Aptitude Test – Verbal Reasoning* (DAT-VR) includes 20 analogy items where participants select a word pair to complete a sentence from five alternatives (e.g., "... is to water as eating is to ..."). The *Differential Aptitude Test – Numerical Reasoning* (DAT-NR) comprises 19 quantitative problems (e.g., solving for a missing digit). Each test was administered within 10 minutes. The *Primary Mental Abilities – Vocabulary* (PMA-V) test consists of 50 synonym items; participants select the word closest in meaning to a target word from four alternatives. The test duration was 4 minutes.

Spatial ability (*Gv*) was assessed using three tasks. The *Rotation of Solid Figures* test includes 21 items in which participants identify a rotated version of a 3D model among four distractors). The *Primary Mental Abilities – Spatial* (PMA-S) test presents a model figure and six alternatives; participants select all correct rotated versions (not mirrored). The score reflects correct responses minus errors (max score = 54); duration was 6 minutes. The *Differential Aptitude Test – Space Relations* (DAT-SR) contains 20 mental folding items where participants match an unfolded shape to its folded version. The time limit was 10 minutes.

Factor Analysis

A confirmatory factor analyses (CFA) was computed using AMOS 16.0.1 (Arbuckle, 2007) for testing the postulated measurement model: three primary factors (*Gf, Gc* and *Gv*) defined by their three intelligence tests, and a higher-order factor representing general intelligence (*g*). The scores for primary factors and for *g* were obtained from the AMOS program. Model fit was checked with the following indices: CMIN/DF, RMSEA, CFI, and SRMR.

MRI data collection

MRIs were obtained with a 3T scanner (GEHC Waukesha, WI, 3T Excite HDX) 8-channels coil. 3D: FSPGR with IR preparation pulse (TR 5.7 ms, TE 2.4 ms TI 750 ms, flip angle 12). Sagital acquisition 0.8 mm thickness, full brain coverage (220 slices), matrix 266×266 FOV 24 (isotropic voxels .7 cm3).

VBM analysis

Voxel-based Morphometry (VBM) was applied for identifying brain areas where GM volumes were correlated with intelligence (Burgaleta et al., 2012; Colom et al., 2009). We used Statistical Parametric Mapping software (SPM8; Ashburner et al., 2012) for pre-processing and statistical analyses. Pre-processing involves image intensity bias correction, segmentation, and normalization. Structural data are divided into different tissue classes using the automated unified segmentation approach provided by the software (Ashburner & Friston, 2005). The modulated GM partitions were then smoothed with a 12-mm FWHM isotropic Gaussian kernel to account for slight misalignments of homologous anatomical structures and to ensure statistical validity under parametric assumptions. Each individual scan was finally fitted to a standardized SPM template specifically created for 3T MRI scans (tissue probability map provided by the International Consortium for Brain Mapping, T1452 Atlas, John C. Mazziotta and Arthur W. Toga, http://www.loni.ucla.edu/Atlases/Atlas_Detail.jsp?atlas_id=6).

The basic design matrix for the statistical analyses was one sample *t* test controlling for sex, age, and handedness ($p < .001$, uncorrected).

Surface-based morphometry

MRI images were also processed by the CIVET pipeline (version 1.1.9) developed at the MNI for fully automated structural image analysis (Ad-Dab'bagh et al., 2006; Kim et al., 2005; MacDonald et al., 2000). CIVET implements a surface-based technique for estimating cortical thickness (CT) and cortical surface area (CSA). Cortical surface area is related to the number and spacing of mini-columnar units of cells in the cerebral cortex, whereas cortical thickness evaluates the number of neurons per column, or neuron density, as well as glial support and dendritic connections (Chance et al., 2008; Lyttelton et al., 2009). Specific stages for the analyses involve (1) registration to MNI-Talairach space (Talairach & Tournoux, 1988), (2) generation of high-resolution hemispheric surfaces with 40962 vertices each, (3) registration of surfaces to a high-resolution template, (4) application of a reverse of step 'a' allowing surface or thickness estimations in native space for each subject, (5) smoothing using a 20-milimeter kernel. See Karama et al. (2009) for further details. Statistical analyses were computed using SurfStat (http://www.math.mcgill.ca/keith/surfstat/) created for MATLAB 7 (The Math-Works, Inc.). Statistical design was based on a *t* test controlling for sex, age, and handedness ($p < .005$, uncorrected).

Results

Intelligence factors

Table 1 shows the correlations among tests along with the descriptive statistics and reliability indices (Cronbach's α) for the nine intelligence measures, primary factors, and *g*.

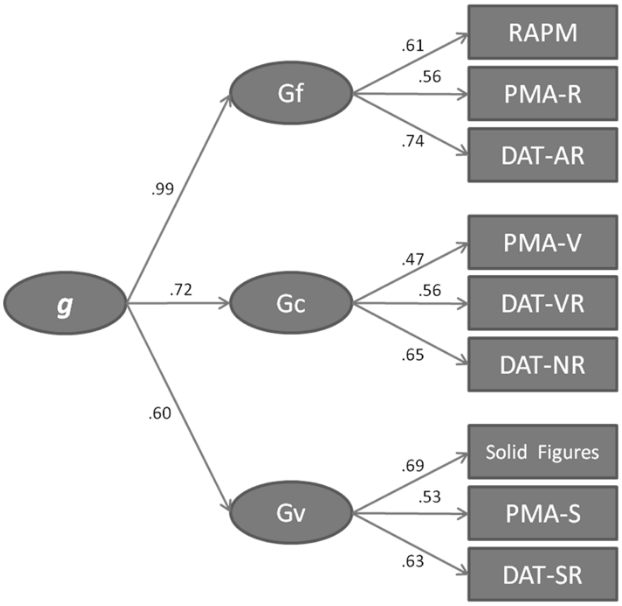
The confirmatory factor analysis (CFA) shows a very good fit: CMIN/DF = .97 (values less than 2.0 denote good fit), RMSEA = .00 (≤ .06 is considered a good fit; Hu & Bentler, 1999), CFI = 1.00 (≥ .90 is considered a good fit) and SRMR = .052 (≤ .08 is considered a good fit; Hu & Bentler, 1999). Figure 1 depicts the structural weights.

Table 1
Descriptive statistics and correlation matrix for the intelligence measures

	1	2	3	4	5	6	7	8	9	10	11	12	13
1. RAPM *(Gf)*		.398**	.467**	.146	.198*	.284**	.236*	.079	.223*	.717**	.509**	.393**	.597**
2. PMA-R *(Gf)*			.371**	.237*	.286**	.316**	.130	.108	.275**	.648**	.538**	.353**	.568**
3. DAT-AR *(Gf)*				.208*	.310**	.345**	.346**	.193*	.371**	.824**	.614**	.550**	.733**
4. PMA-V *(Gc)*					.296**	.311**	.201*	.217*	-.010	.348**	.560**	.249*	.427**
5. DAT-VR *(Gc)*						.343**	.114	.081	.241*	.466**	.649**	.270**	.511**
6. DAT-NR *(Gc)*							.200*	.113	.193	.550**	.792**	.337**	.619**
7. Solid Figures *(Gv)*								.415**	.408**	.453**	.358**	.833**	.606**
8. PMA-S *(Gv)*									.322**	.293**	.252*	.643**	.438**
9. DAT-SR *(Gv)*										.474**	.333**	.725**	.565**
10. Fluid intelligence *(Gf)*											.863**	.721**	.953**
11. Crystallized intelligence *(Gc)*												.590**	.905**
12. Spatial Intelligence *(Gv)*													.853**
13. *g* factor													
Mean	11.84	11.97	14.43	32.69	13.64	11.95	9.00	27.53	15.95				
SD	2.36	4,49	3.52	6.57	3.00	3.23	3.89	9.67	4.79				
a	.66	.87	.88	.79	.68	.82	.74	.73	.84				

Note. *p <. 05; **p < .01). *Raven Advanced Progressive Matrices test (RAPM), abstract reasoning (DAT-AR), verbal reasoning (DAT-VR), spatial relations (DAT-SR) and numerical reasoning (DAT-NR) subtests from the Differential Aptitude test (DAT-5) Battery. Rotation of Solid Figures (Solid Figures) and from the Primary Mental Abilities Battery (PMA) the subtest: inductive reasoning (PMA-R), vocabulary subtest (PMA-V) and spatial subtest (PMA-S) and primary factors: fluid intelligence (Gf), verbal-crystallized intelligence (Gc) and spatial intelligence (Gv) and higher-order factor or general intelligence (g).*

Figure 1
Confirmatory model for the considered measures of intelligence.



Note. RAPM = Raven Advanced Progressive Matrices Test, PMA-R = inductive reasoning subtests, DAT-AR = abstract reasoning subtest, PMA-V = vocabulary subtests, DAT-VR = verbal reasoning subtest, DAT-NR = numerical reasoning subtest test, PMA-S = mental rotation subtest, DAT-SR = spatial relations subtest and Solid Figures = Rotation of Solid Figures), the three primary factors (Gf = fluid intelligence, Gc = verbal- crystallized intelligence and Gv = spatial intelligence), and the higher-order factor (g).

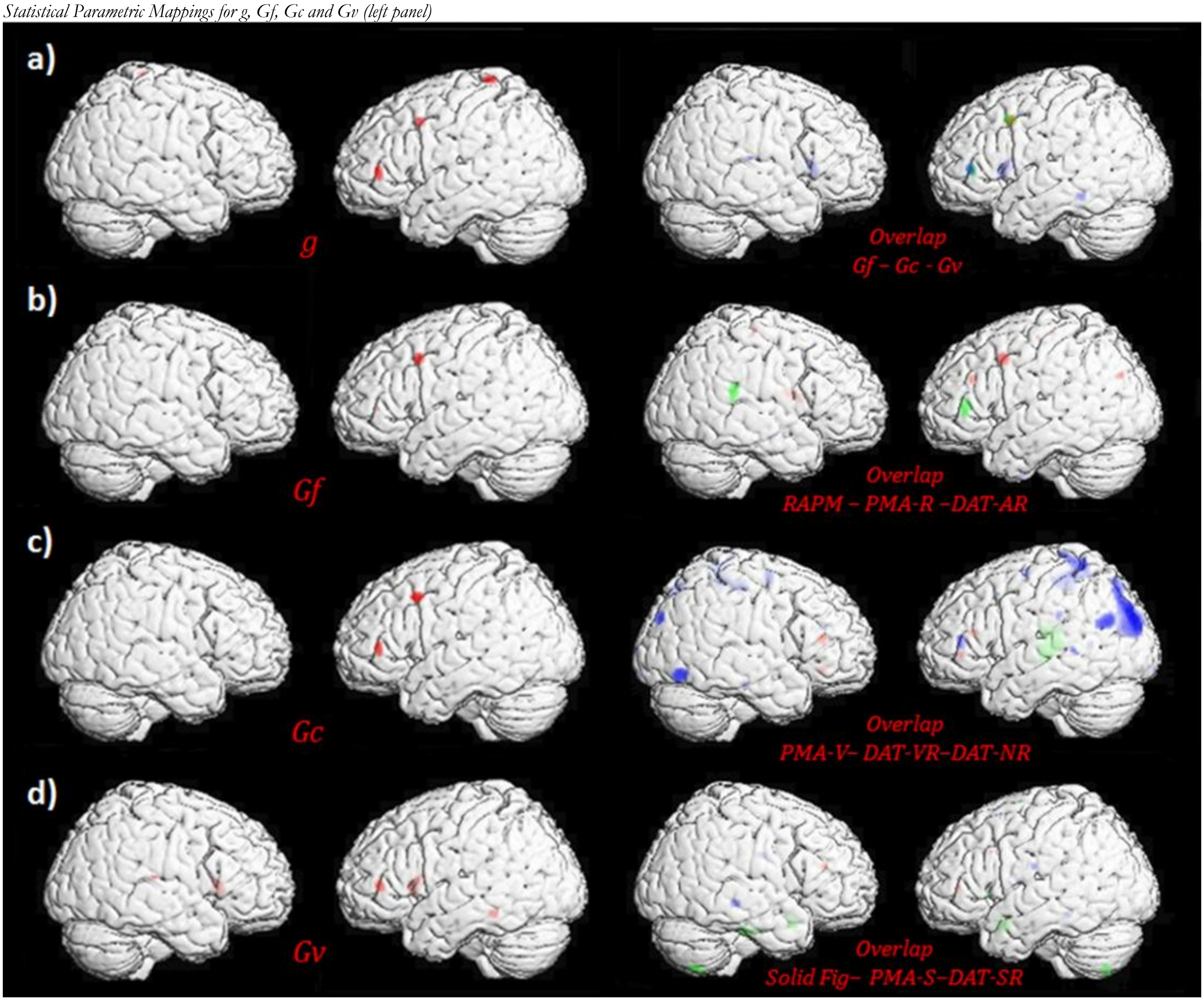
Voxel- based Morphometry

The Appendix reports the peak voxel coordinates of brain clusters showing significant correlations between regional gray matter volume and different levels of the intelligence hierarchy: (a) measures, (b) first-order factors (*Gf, Gc*, and *Gv*), and (c) the higher-order factor (*g*). As noted above, results are shown using a *p* level of .001 uncorrected for multiple comparisons. Gray matter correlations were found at the test level, for the group factors, and for *g*. All regions for *g* are located in Right-Frontal regions, and the same pattern emerges for fluid-abstract intelligence (Gf) and crystallized-verbal (Gc) intelligence. Frontal, temporal, and subcortical regions are correlated with spatial intelligence. Note that we also computed analyses for the RAPM, Gf, and *g* excluding left-handed participants and results did not change.

Overlapping and non-overlapping areas

SPMs for the intelligence latent factors and tests were overlapped using the display option provided by the software. Figure 2 shows the significant areas for (a) the *g* factor and the three group factors; (b) fluid intelligence (*Gf*) and its three tests; (c) verbal-crystallized intelligence (*Gc*) and its tests, and (d) spatial intelligence (*Gv*) and its tests. All these findings are also shown using a *p* level of .001 uncorrected for multiple comparisons, controlling for sex, age and handedness (the specific coordinates are reported in the Appendix).

Figure 2
Statistical Parametric Mappings for g, Gf, Gc and Gv (left panel)



Note. The right panel shows the overlap for (a) fluid intelligence (Gf – red), verbal-crystallized intelligence (Gc – green), and spatial intelligence (Gv – blue); (b) Raven Advanced Progressive Matrix (RAPM-red), inductive reasoning (PMA-R - green) and abstract reasoning (DAT-AR - blue); (c) vocabulary (PMA-V - red), verbal reasoning (DAT-VR - green) and numerical reasoning (DAT-NR - blue), and finally (d) Rotation of Solid Figures (Solid Fig. - red), spatial relations (DAT-SR - green) and mental rotation (PMA-S - blue). p < .001 (uncorrected for multiple comparisons) controlling for Sex, Age, and Handedness.

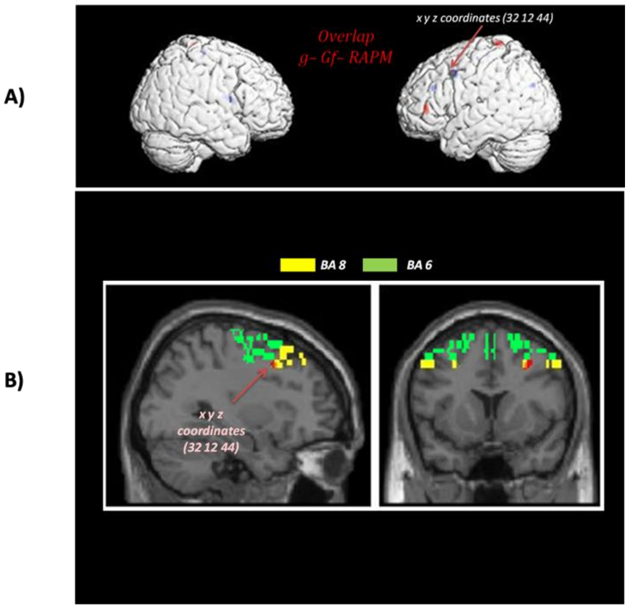
Overlap is shown in BA 10 (MNI coordinates 34, 42, 6) for *g*, *Gc* and *Gv*. DAT-NR overlaps with *Gc* and *g*, but not on the same coordinates. Still another overlap is observed in right BA 8 (MNI coordinates 32, 12, 44) for the *g* factor, fluid intelligence (*Gf*), and the Raven Advanced Progressive Matrix Test (RAPM). Figure 3-A highlights this finding.

VBM results were obtained using an uncorrected threshold of $p < .001$ at the whole-brain level. As required, we checked if this key region survives correction for multiple comparisons. Importantly, as discussed by Salgado-Pineda et al. (2003) corrections over the whole brain are very strict when applied to structural data. Hence, this analysis was restricted to voxels contained in the (very big) frontal lobe following Gong et al. (2005). This was done using the small volume correction (SVC) function in SPM-8 (FWE, $p < .05$). Using a sphere of 20x20x20 centered in BA 9 (Colom et al., 2009; Gong et al., 2005; Haier et al., 2004; Jung & Haier, 2007) the region passed the test.

Therefore, this latter result is consistent with the hypothesis that gray matter correlates for the RAPM are closely similar to *Gf*/*g*. This happens even when the RAPM is not the intelligence measure with the highest correlation with *g* at the behavioral level (see Table 1). The key brain cluster shared by statistical parametric mappings for *g*, *Gf*, and the RAPM, approximately involves Brodmann area 8 (BA 8). The topographic location of the identified cluster suggests Brodmann area 6 (BA 6) may be also implicated (see Figure 3-B).

Figure 3

A) Overlap among g (red), Gf (green), and RAPM (blue), B) Cluster found in VBM analyses, and its localization in the junction between BAs 6 and 8



Note. *$p < .001$ (uncorrected for multiple comparisons) controlling for Sex, Age, and Handedness. The specific overlap for the three levels of intelligence hierarchy is highlighted.*

Surface-based morphometry results

We analyzed the same dataset using a different morphometry approach, as noted above, finding significant results for cortical surface area, but not for cortical thickness.

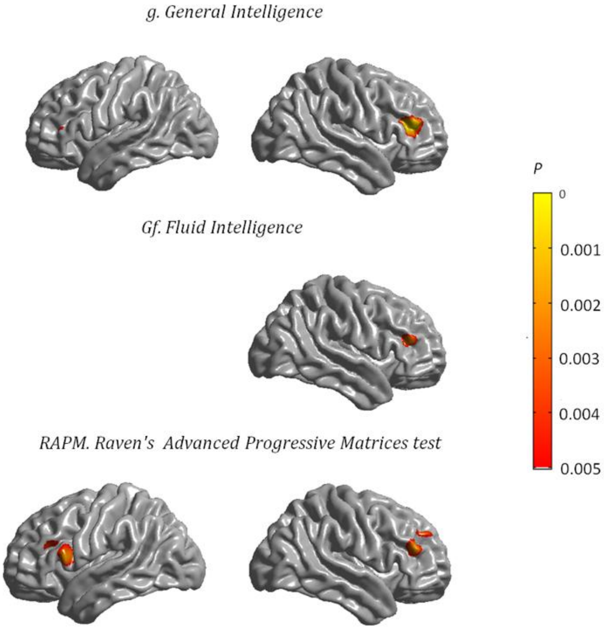
Figure 4 shows results for cortical surface area in Right BA 46 (MNI coordinates 44, 37, 14). This region is significant for the three levels of intelligence hierarchy: *g* – *Gf* – RAPM ($p < .005$, controlling for sex, age and handedness). The RAPM is again the only test showing a very specific overlap with both the higher-order *g* factor and the primary factor Gf. Nevertheless, still other significant areas were found for *g* (Left BA 46) and for the RAPM (Right BA 9, Left BA 44, 46). SBM analyses were conducted using vertex-wise statistics uncorrected for multiple comparisons at the whole-brain level; however, the significant BA 46 cluster survives to Small Volume Correction (SVC, FWE, p < .05).

Discussion

Here we have reported gray matter correlates for psychological scores representing different levels of the intelligence hierarchy (Colom et al., 2009; Colom & Thompson, 2011; Haier et al., 2009a, 2023; Karama et al., 2011). For the intelligence tests, notable widespread results were observed, which is consistent with Haier et al. (2010). This also emerges for the first-order group factors representing fluid-abstract, crystallized-verbal, and spatial intelligence (see Román et al., 2014 for a full discussion). Results were obtained using an uncorrected threshold of $p < .001$ at the whole-brain level; therefore, these findings should be interpreted as exploratory, with the strongest inferences drawn from the SVC-corrected results. After applying SVC correction, a specific overlap emerged across the three levels of the intelligence hierarchy— Raven's Advanced Progressive Matrices Test (RAPM), *Gf*, and the general factor (*g*)—highlighting the robustness of this shared neural substrate. This finding supports the special status of the RAPM as one specific measure tapping core facets of the intelligence construct. Therefore, the main prediction is supported.

Figure 4

Map of p-values ($p < .005$) for the relationship between g, Gf, RAPM scores and cortical surface area



Note. Significant correlations between cortical surface area and (a) g (general intelligence), (b) Gf (fluid intelligence) and (c) RAPM (Advanced Progressive Matrices Test).

Regarding the main outcomes of the present study, the key brain cluster common to the statistical parametric maps for general intelligence (g), fluid intelligence (Gf), and Raven's Advanced Progressive Matrices (RAPM) includes Brodmann areas 8 (BA 8), 6 (BA 6) and 46 (BA 46). It is worth noting, as highlighted by Lashley and Clark (1946) several decades ago "there is significant individual variation in the architectonic structure of areas similarly located in different specimens of the same species" (cited by Zilles & Amunts, 2010).

BA 8 supports planning of complex movements. There are some previous findings for this area regarding intelligence (Colom et al., 2006a, 2006b; Gong et al., 2005). Surface-based correlates at BA 8 have been reported in children and adolescents (Karama et al., 2009; Shaw et al., 2006). BA 8 is also related to working memory capacity. Thus, for instance, du Boisgueheneuc et al. (2006) compared control and lesion patients in the n-back task, finding that a lesion in BA 8 significantly impairs performance. Further, Rowe and

Passingham (2001) found that BA 8 is involved in the maintenance of spatial locations in working memory.

The Ramsden et al.'s (2011) study showed changes in verbal and non-verbal IQ during teenage years (from 14 to 18 years) associated with changes in brain regions not directly related with areas underscored by the P-FIT model. Specifically, changes in verbal IQ were associated with motor speech region (BA 6) whereas non-verbal IQ changes were related to regions associated with finger movements (Anterior Cerebellum). These changes in IQ related to changes in gray matter show that P-FIT areas might also include these regions for capturing the dynamic nature of the intelligence construct. After all, this model may be seen as a dynamic framework open to new evidence (Haier et al., 2023). In the Ramsden et al. study, changes in non-verbal IQ (closely similar to *Gf*) were related to regions associated with body movements (BA 6) which is consistent with the key finding reported here (see Burgaleta et al., 2014).

In addition, functional studies using PET revealed that a measure of frontal lobe function and visuo-spatial reasoning (perceptual maze task) activates BA 8 (Ghatan et al., 1995). Goel et al. (1997) reported that inductive reasoning is also associated with BA 8. Duncan et al. (2000) found that highly *g* loaded tasks evoke activations spread across frontal regions including BAs 6 and 8. On the other hand, *f*MRI studies report results for BA 8 and performance in chess (Atherton et al., 2003) or fluid intelligence measures (Fangmeier et al., 2006; Geake & Hansen, 2005). This region is also highlighted in studies reporting changes in structural and functional MRI after training in complex tasks (Haier et al., 2009b).

Colom et al. (2009) proposed that BA 8 might work in tandem with other frontal areas towards evaluation and hypothesis testing components. Only three clusters belonging to BAs 4, 8, and 10 were detected for *g*, and only findings for the RAPM overlapped one of these clusters. Nevertheless, group factors showed overlap in some areas: fluid and verbal-crystallized intelligence (BA 8), as well as verbal-crystallized and spatial intelligence (BA 10). This is noteworthy because neuroimaging findings rarely overlap with such specificity (Colom, 2007; Hunt, 2011; see Basten et al., 2015).

Crucially, findings derived from the analysis of cortical surface area (CSA) indices using a different morphometry technique (SBM) support the main conclusion. The Progressive Matrices Test is the only specific measure overlapping with both *Gf* and *g*. Here, the relevant overlapping cluster belongs to BA 46 (dorsolateral prefrontal cortex, DLPFC; corrected by SVC) a brain region supporting working memory capacity and highlighted by the parieto-frontal integration theory (P-FIT) of intelligence (Jung & Haier, 2007). Interestingly, significant associations were observed for cortical surface area (CSA) but not for cortical thickness (CT). This dissociation provides valuable insight into the neurobiological mechanisms underlying the relationship between intelligence and brain structure. While CSA is thought to reflect variations in the number and spacing of cortical columns established during early neurodevelopment, CT is more closely associated with neuronal density, dendritic arborization, and cortical layering (Panizzon et al., 2009). The present findings therefore suggest that the association between RAPM and *g* may be primarily driven by large-scale surface area expansion rather than by differences in cortical thickness. This interpretation aligns with evidence linking CSA to global cognitive ability and individual differences in intellectual performance, highlighting the role of early neurodevelopmental processes in shaping the neural architecture that supports higher-order cognition (Vuoksimaa et al., 2016).

BA 46 is considered a hub in brain networks (Sporns et al., 2007), which serves a connecting functional role in large-scale cortical modeling (Honey et al., 2007; Hilger & Sporn, 2021). This region receives sensory inputs from the posterior cortex and integrates external information with internal goals, keeping this information for action (Goldman-Rakic, 1996; Quintana & Fuster, 1999). Also, BA 46 is involved in working memory, spatial and verbal cognition, and selective attention (Kane & Engle, 2002; Ramnani & Owen, 2004; Rowe & Passingham, 2001; Wager & Smith, 2003). Furthermore, this region is related to intelligence (Colom et al., 2006a, 2009; Haier et al., 2004, 2009a). It is important to underscore that these regions (BA 8, BA 46, and BA 6) are located in the dorsolateral prefrontal cortex, which is critical for monitoring events in working memory (Petrides, 2000). Working memory and intelligence are highly related (Ackerman et al., 2005; Colom et al., 2016; Kane et al., 2004; Martínez et al., 2011).

If the RAPM is the best single estimate of general intelligence (*g*), and *g* is strongly related with working memory, it is not surprising that regions found for the RAPM are also involved in working memory. RAPM performance likely depends not only on working memory capacity but also on the efficient coordination of short-term memory and higher-order reasoning processes, reflecting a shared neural substrate for complex cognition. This interpretation aligns with findings showing that fluid intelligence is closely linked to short-term storage, maintenance, and updating processes (Martínez et al., 2011; Wilhelm et al., 2013).

Although it has been argued that the RAPM primarily reflects lower-level perceptual or visuospatial processes (Spearman & Wynn-Jones, 1951), our findings show convergence in high-level prefrontal regions (BA 8 and BA 46). These areas are associated with abstract reasoning and working memory, providing neural evidence against the low-level perceptual interpretation and reinforcing the view that RAPM engages higher-order cognitive mechanisms. This perspective aligns with prior work showing that the reasoning processes required by the RAPM recruits dorsolateral prefrontal and parietal areas rather than purely visuospatial systems (Gray et al., 2003; Duncan et al., 2000). Critically, behavioural research has specifically highlighted concerns about the visuo spatial bias of RAPM (Gonthier, 2022). Thus, our structural neuroimaging evidence contributes to this debate by showing that RAPM's neural correlates align more closely with executive and reasoning networks than mainly with perceptual or visuospatial systems.

A limitation of the present study concerns the restricted age and educational range of the sample, composed primarily of young university students. This homogeneity may reduce interindividual variability in cognitive and neural measures, potentially limiting generalizability to other age or educational groups. However, it also minimizes confounding effects related to brain maturation and cognitive performance, allowing for a more precise examination of the structural correlates of intelligence within a developmentally stable sample. In addition, although the RAPM short form has a modest internal consistency (α = .66), the robustness of the neuroimaging findings suggests a strong underlying relationship between RAPM performance and the neural substrates of general intelligence. This indicates that, despite the limitations of the short form, the test effectively captures key aspects of the cognitive processes reflected in the brain regions identified, reinforcing its validity as an index of *g*.

In conclusion, the reported results are consistent with the statement that the Raven Progressive Matrices Test is the best available single estimate of the general factor of intelligence (*g*), as underscored by several reports (Deary & Stough, 1996; Deshon et al., 1995; Jensen, 1998; McGrew & Flanagan, 1998; Spearman, 1938; Spearman & Wynn-Jones, 1951). Although psychometric and cognitive studies raise reasonable doubts with respect to this general conclusion, the exhaustive analysis relating the intelligence hierarchy and gray matter correlates for scores derived from the specific measures, first-order group factors, and the general factor of intelligence (*g*) suggests that the matrices test is the best choice in neuroimaging studies for estimating the general factor of intelligence (*g*). Novel research combining different neuroimaging approaches for decoding the brain during intelligence testing reinforces this conclusion (Thiele et al., 2025).

Complementary information

Conflict of interest: The authors declare no conflict of interest.

Financial support: This research was partially supported by Grant PSI2010-20364 (*Ministerio de Ciencia e Innovación*, Spain) and PID2022-141787NB-I00 (*Ministerio de Ciencia, Innovación y Universidades*, Spain). FJR and MB were supported by PID2022-141787NB-I00 (*Ministerio de Ciencia, Innovación y Universidades*, Spain).

Data Research Availability Statement: The data used for this study are available upon reasonable request to the corresponding author.

References

Abad, F. J., Colom, R., Rebollo, I., & Escorial, S. (2004). Sex differential item functioning in the Raven's Advanced Progressive Matrices: Evidence for bias. *Personality and Individual Differences, 36,* 1459−1470. http://dx.doi.org/10.1016/S0191-8869(03)00241-1

Ackerman, P. L. (1999). Traits and knowledge as determinants of learning and individual differences: Putting it all together. In P. L. Ackerman, P. C. Kyllonen, & R. D. Roberts (Eds.), *Learning and individual differences: Process, trait, and content determinants* (pp. 437–460). American Psychological Association.

Ackerman, P. L., Beier, M. E., & Boyle, M. O. (2002). Individual Differences in Working Memory within a Nomological Network of Cognitive and Perceptual Speed Abilities. *Journal of Experimental Psychology: General, 131,* 567 -589. http://dx.doi.org/10.1037/0096-3445.131.4.567

Ackerman, P. L., Beier, M. E., & Boyle, M. O. (2005). Working memory and intelligence: The same or different constructs? *Psychological Bulletin, 131,* 30−60. http://dx.doi.org/10.1037/0033-2909.131.1.72

Ad-Dab'bagh, Y, Lyttelton, O, Muehlboeck, J-S, Lepage, C, Einarson, D, Mok, K, Ivanov, O, Vincent, RD, Lerch, J, Fombonne, E, Evans, AC (2006). The CIVET image processing environment: a fully automated comprehensive pipeline for anatomical neuroimaging research. In: Corbetta, M (Ed), *Proceedings of the 12th Annual Meeting of the Organization for Human Brain Mapping NeuroImage*, Florence, Italy.

Alderton, D. L., & Larson, G. E. (1990). Dimensionality of Raven's Advanced Progressive Matrices. *Educational and Psychological Measurement, 50,* 887–900. http://dx.doi.org/10.1177/0013164490504019

Arbuckle, J. L. (2007). *AMOS, Version 16.0.1.* Amos Development Corporation.

Arthur, W., & Woehr, D. J. (1993). A confirmatory factor analytic study examining the dimensionality of the Raven's Advanced Progressive Matrices. *Educational and Psychological Measurement, 53*, 471–497. http://dx.doi.org/10.1177/0013164493053002016

Ashburner, J., & Friston, K. J. (2005). Unified segmentation. *NeuroImage, 26,* 839–851. http://dx.doi.org/10.1016/j.neuroimage.2005.02.018

Ashburner, J., Barnes, G., Chen, C., Daunizeau, J., Flandin, G., Friston, K., ... & Phillips, C. (2012). SPM8 manual. *Functional Imaging Laboratory, Institute of Neurology, 41.*

Atherton, M., Zhuang, J., Bart, W. M., Hu, X. & He, S. (2003). A functional MRI study of high-level cognition. I. The game of chess. *Cognitive Brain Research, 16,* 26–31. http://dx.doi.org/10.1016/S0926-6410(02)00207-0

Basten, U., Hilger, K., & Fiebach, C. J. (2015). Where smart brains are different: A quantitative meta-analysis of functional and structural brain imaging studies on intelligence. *Intelligence*, *51*, 10-27. https://doi.org/10.1016/j.intell.2015.04.009

Bennett, G. K., Seashore, H. G., & Wesman, A. G. (1990). *Differential Aptitude Test (5th Edition).* TEA.

Burgaleta, M., Head, K., Álvarez-Linera, J., Martínez, K., Escorial, S., Haier, R., & Colom, R. (2012). Sex differences in brain volume are related to specific skills, not to general intelligence. *Intelligence, 40,* 60–68. http://dx.doi.org/10.1016/j.intell.2011.10.006

Burgaleta, M., Johnson, W., Waber, D. P., Colom, R., & Karama, S. (2014). Cognitive ability changes and dynamics of cortical thickness development in healthy children and adolescents. *Neuroimage*, *84*, 810-819. https://doi.org/10.1016/j.neuroimage.2013.09.038

Burke, H. R. (1958). Raven's Progressive Matrices: A review and critical evaluation. *Journal of Genetic Psychology, 93,* 199–228.

Carpenter, P. A., Just, M. A., & Shell, P. (1990). What one intelligence test measures: A theoretical account of the processing in the Raven Progressive Matrices Test. *Psychological Review, 97*, 404–431. http://dx.doi.org/10.1037/0033-295X.97.3.404

Carroll, J. B. (1993). *Human cognitive abilities.* Cambridge University Press.

Cattell, R. B. (1980). The heritability of fluid, gf, and crystallised, gc, intelligence, estimated by a least squares use of the MAVA method. *British Journal of Educational Psychology, 50,* 253-265.

Chance, S. A., Casanova, M. F., Switala, A. E. & Crow, T. J. (2008). Auditory cortex asymmetry, altered minicolumn spacing and absence of ageing effects in schizophrenia. *Brain, 131*, 3178–3192. http://dx.doi.org/10.1093/brain/awn211

Colom, R. (2007). Intelligence? What Intelligence? *Behavioural and Brain Sciences, 30*, 155-156. http://dx.doi.org/10.1017/S0140525X07001215

Colom, R., & Abad, F. J. (2007). Advanced progressive matrices and sex differences: Comment to Mackintosh and Bennett (2005). *Intelligence, 35*, 183−185.

http://dx.doi.org/10.1016/j.intell.2006.06.003http://dx.doi.org/10.1016/j.intell.2008.01.002

Colom, R., Chuderski, A., & Santarnecchi, E. (2016). Bridge over troubled water: Commenting on Kovacs and Conway's process overlap theory. *Psychological Inquiry, 27*(3), 181–189. https://doi.org/10.1080/1047840X.2016.1181513

Colom, R., Escorial, S., & Rebollo, I. (2004). Sex differences on the progressive matrices are influenced by sex differences on spatial ability. *Personality and Individual Differences, 37,* 1289–1293. http://dx.doi.org/10.1016/j.paid.2003.12.014

Colom, R., & Garcia-Lopez, O. (2002). Sex differences in fluid intelligence among high school graduates. *Personality and Individual Differences, 32,* 445–451. http://dx.doi.org/10.1016/S0191-8869(01)00040-X

Colom, R., Haier, R. J., Head, K., Álvarez-Linera, J., Quiroga, M. A., Shih, P. C., & Jung, R. E. (2009). Gray matter correlates of fluid, crystallized, and spatial intelligence: Testing the P-FIT model. *Intelligence, 37,* 124–135. http://dx.doi.org/10.1016/j.intell.2008.07.007

Colom, R., Jung, R. E. & Haier, R. J. (2006a). Finding the g-factor in brain structure using the method of correlated vectors. *Intelligence, 34*, 561–570. http://dx.doi.org/10.1016/j.intell.2006.03.006

Colom, R., Jung, R. E. & Haier, R. J. (2006b). Distributed brain sites for the g-factor of intelligence. *NeuroImage, 31,* 1359–1365. http://dx.doi.org/10.1016/j.neuroimage.2006.01.006

Colom, R., Karama, S., Jung, R. E., & Haier, R. J. (2010). Human intelligence and brain networks. *Dialogues in Clinical Neuroscience, 12,* 489-501.

Colom, R. & Thompson, P.M. (2011). Understanding Human Intelligence by imaging the Brain. In T. Chamorro-Premuzic, A. Furnham, & S. von Stumm (Eds.), *Handbook of individual differences*. Wiley-Blackwell.

Deary, I. J., & Stough, C. (1996). Intelligence and inspection time: Achievements, prospects, and problems. *American Psychologist, 51,* 599–608. http://dx.doi.org/10.1037/0003-066X.51.6.599

Deshon, R. P., Chan, D., & Weissbein, D. A. (1995). Verbal overshadowing effects on Raven's Advanced Progressive Matrices: evidence for multidimensional performance determinants. *Intelligence, 21,* 135–155. http://dx.doi.org/10.1016/0160-2896(95)90023-3

Dillon, R. F., Pohlmann, J. T., & Lohman, D. F. (1981). A factor analysis of Raven's Advanced Progressive Matrices. *Educational and Psychological Measurement, 41,* 1295–1302. http://dx.doi.org/10.1177/001316448104100438

Dolke, Ashok M. (1976). Investigation into certain psychometric properties of Raven's Standard Progressive Matrices Test. *Indian Journal of Psychology, 51,* 225-236.

du Boisgueheneuc, F., Levy, R., Volle, E., Seassau, M., Duffau, H., Kinkingnehun, S., Samson, Y., Zhang, S & Dubois, B. (2006). Functions of the left superior frontal gyrus in humans: a lesion study. *Brain, 129,* 3315–3328. http://dx.doi.org/10.1093/brain/awl244

Duncan, J., Seitz, R.J., Kolodny, J., Bor, D., Herzog, H., Ahmed, A., Newell, F.N., Emslie, H. (2000). A neural basis for general intelligence. *Science, 289*, 457– 460. http://dx.doi.org/10.1126/science.289.5478.457

Fangmeier, T., Knauff, M., Ruff, C. C. & Sloutsky, V. (2006). fMRI evidence for a three-stage model of deductive reasoning. *Journal of Cognitive Neuroscience, 18,* 320–334. http://dx.doi.org/10.1162/jocn.2006.18.3.320

Ghatan, P. H., Hsieh, J. C., Wirsen-Meurling, A., Wredling, R., Eriksson, L., Stone-Elander, S., Levander, S. & Ingvar, M. (1995). Brain activation induced by the perceptual maze test: A PET study of cognitive performance. *NeuroImage, 2,* 112–124. http://dx.doi.org/10.1006/nimg.1995.1014

Gonthier, C. (2022). Cross-cultural differences in visuo-spatial processing and the culture-fairness of visuo-spatial intelligence tests: An integrative review and a model for matrices tasks. *Cognitive Research: Principles and Implications*, *7*, 11. https://doi.org/10.1186/s41235-021-00350-w

Geake, J. G. & Hansen, P. C. (2005). Neural correlates of intelligence as revealed by fMRI of fluid analogies. *NeuroImage, 26,* 555–564. http://dx.doi.org/10.1016/j.neuroimage.2005.01.035

Goel, V., Gold, B., Kapur, S. & Houle, S. (1997). The seats of reason? An imaging study of deductive and inductive reasoning. *NeuroReport, 8,* 1305– 1310.

Goldman-Rakic, P. S. (1996) The prefrontal landscape: implications of functional architecture for understanding human mentation and the central executive. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences, 351*, 1445–1453.

Gong, Q. Y., Sluming, V., Mayes, A., Keller, S., Barrick, T., Cezayirli, E. & Roberts, N. (2005). Voxel-based morphometry and stereology provide convergent evidence of the importance of medial prefrontal cortex for fluid intelligence in healthy adults. *NeuroImage, 25,* 1175–86. http://dx.doi.org/10.1016/j.neuroimage.2004.12.044

Gray, J. R., Chabris, C. F., & Braver, T. S. (2003). Neural mechanisms of general fluid intelligence. *Nature Neuroscience, 6,* 316–322. http://dx.doi.org/10.1038/nn1014

Haier, R. J., Colom, R., Schroeder, D.H., Condon, C.A., Tang, C, Eaves, E. & Head, K. (2009a). Gray matter and intelligence factors: Is there a neuro-g? *Intelligence, 37,* 136-134. http://dx.doi.org/10.1016/j.intell.2008.10.011

Haier, R. J., Colom, R., Hunt, E. B. (2023). *The science of human intelligence.* Cambridge University Press.

Haier, R. J., Jung, R, E., Yeo, R. A., Head, K., & Alkired, M. T. (2004). Structural brain variation and general intelligence. *Neuroimage, 23,* 425 – 433. http://dx.doi.org/10.1016/j.neuroimage.2004.04.025

Haier, R. J., Karama, S., Leyba, L., & Jung, R. E. (2009b). MRI assessment of cortical thickness and functional activity changes in adolescent girls following three months of practice on a visual-spatial task. *BMC Research Notes, 2,* 174. http://dx.doi.org/10.1186/1756-0500-2-174

Haier, R. J., Schroeder, D., Tang, C., Head, K., & Colom, R. (2010). Gray matter correlates of cognitive ability tests used for vocational guidance. *BMC Research Notes.* http://www.biomedcentral.com/1756-0500/3/206

Haier, R. J., Siegel, B. V., Nuechterlein, K. H., Hazlett, E., Wu, J. C., Paek, J., Browning, H. L., & Buchsbaum, M. S. (1988). Cortical glucose metabolic rate correlates of abstract reasoning and attention studied with positron emission tomography. *Intelligence, 12,* 199-217. http://dx.doi.org/10.1016/0160-2896(88)90016-5

Haier, R. J., White, N. S., Alkire, M. T. (2003). Individual differences in general intelligence correlate with brain function during nonreasoning tasks. *Intelligence, 31,* 429–441. http://dx.doi.org/10.1016/S0160-2896(03)00025-4

Hilger, K., & Sporns, O. (2021). Network neuroscience methods for studying intelligence. In A. K. Barbey, S. Karama, & R. J. Haier (Eds.), *The Cambridge handbook of intelligence and cognitive neuroscience* (pp. 26–43). Cambridge University Press. https://doi.org/10.1017/9781108635462.004

Honey, C. J., Kötter, R., Breakspear, M., & Sporns, O. (2007) Network structure of cerebral cortex shapes functional connectivity on multiple time scales. *Proceedings of the National Academy of Sciences, 104,* 10240–10245. http://dx.doi.org/10.1073/pnas.0701519104

Horn, J. (1985). Remodeling old models of intelligence. In B. B. Wolman (Ed.), *Handbook of intelligence.* Wiley.

Hu, L. T. & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structural analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6,* 1 – 55. http://dx.doi.org/10.1080/10705519909540118

Hunt, E. B. (1974). Quote the Raven? Nevermore! In L. W. Gregg (Ed.), *Knowledge and cognition.* Erlbaum.

Hunt, E. B. (2011). *Human Intelligence.* Cambridge University Press.

Jensen, A. R. (1980). Author's response. Précis of bias in mental testing. *Behavioral and Brain Sciences, 3,* 359–368.

Jensen, A. R. (1998). *The g factor: The science of mental ability.* Praeger.

Jung, R. E. & Haier, R. J. (2007). The parieto-frontal integration theory (P-FIT) of intelligence: Converging neuroimaging evidence. *Behavioral & Brain Sciences*, *30*, 135–187. http://dx.doi.org/10.1017/S0140525X07001185

Kane, M. J., & Engle, R. W. (2002). The role of prefrontal cortex in working memory capacity, executive attention, and general fluid intelligence: An individual-differences perspective. *Psychonomic Bulletin & Review, 9,* 637−671. http://dx.doi.org/10.3758/BF03196323

Kane, M. J., Hambrick, D. Z., Tuholski, S. W., Wilhelm, O., Payne, T. W., & Engle, R. W. (2004). The generality of working memory capacity: A latent-variable approach to verbal and visuospatial memory span and reasoning. *Journal of Experimental Psychology-General, 133,* 189−217. http://dx.doi.org/10.1037/0096-3445.133.2.189

Karama, S., Ad-Dab'bagh, Y., Haier, R. J., Deary, I. J., Lyttelton, O. C., Lepage, C., Evans, A. C. & the Brain Development Cooperative

Group. (2009). Positive association between cognitive ability and cortical thickness in a representative US sample of healthy 6 to 18 year-olds. *Intelligence, 37,* 145–155. http://dx.doi.org/10.1016/j.intell.2008.09.006

Karama, S., Colom, R., Jhonson, W., Deary, I. J., Haier, R., Waber, D. P., Lepage, C., Ganjavi, H., Jung, R., Evans, A. C., & the Brain Development Cooperative Group. (2011). Cortical thickness correlates of specific cognitive performance accounted for by the general factor of intelligence in healthy children aged 6 to 18. *Neuroimage, 55,* 1443–1453. http://dx.doi.org/10.1016/j.neuroimage.2011.01.016

Kim, J. S., Singh, V, Lee, J. K., Lerch, J., Ad-Dab'bagh, Y., MacDonald, D., Lee, J. M., Kim, S. I., Evans, A. C., (2005). Automated 3-D extraction and evaluation of the inner and outer cortical surfaces using a Laplacian map and partial volume effect classifcation. *NeuroImage, 27*, 210–221. http://dx.doi.org/10.1016/j.neuroimage.2005.03.036

Lashley, K. S., & Clark, G. (1946). The cytoarchitecture of the cerebral cortex of Ateles: a critical examination of architectonic studies. *Journal of Comparative Neurology, 85,* 223–305.

Lee, K. H., Choi, Y. Y., Gray, J. R., Cho, S. H., Chae, J. H., Lee, S. & Kim, K. (2006). Neural correlates of superior intelligence: Stronger recruitment of posterior parietal cortex. *NeuroImage, 29,* 578–586. http://dx.doi.org/10.1016/j.neuroimage.2005.07.036

Lynn, R., Allik, J. & Irwing, P. (2004). Sex differences on three factors identified in Raven's Standard Progressive Matrices. *Intelligence, 32,* 411–424. http://dx.doi.org/10.1016/j.intell.2004.06.007

Lyttelton, O. C., Karama, S., Ad-Dab'baghm, Y., Zatorre, R. J., Carbonell, F., Worsley, K. & Evans, A. (2009). Positional and surface area asymmetry of the human cerebral cortex. *NeuroImage*, *46,* 895-903. http://dx.doi.org/10.1016/j.neuroimage.2009.03.063

Mackintosh, N. J., & Bennett, E. S. (2005). What do Raven's Matrices measure? An analysis in terms of sex differences. *Intelligence, 33,* 663−674. http://dx.doi.org/10.1016/j.intell.2005.03.004

Martínez, K., Burgaleta, M., Román, F. J., Escorial, S., Shih, P. C., Quiroga, Mª A., & Colom, R. (2011). Can Fluid Intelligence Be Reduced To 'Simple' Short-Term Storage? *Intelligence, 39*, 473-480. http://dx.doi.org/10.1016/j.intell.2011.09.001

Martínez, K., & Colom, R. (2021). Imaging the intelligence of humans. In A. K. Barbey, S. Karama, & R. J. Haier (Eds.), *The Cambridge handbook of intelligence and cognitive neuroscience* (pp. 44–69). Cambridge University Press. https://doi.org/10.1017/9781108635462.005

MacDonald, D., Kabani, N., Avis, D., Evans, A. C. (2000). Automated 3-D extraction of inner and outer surfaces of cerebral cortex from MRI. *NeuroImage, 12,* 340–356. http://dx.doi.org/10.1006/nimg.1999.0534

Malkiński, M., & Mańdziuk, J. (2025). Deep learning methods for abstract visual reasoning: A survey on raven's progressive matrices. *ACM Computing Surveys, 57,* 1-36. https://doi.org/10.1145/3715093

McGrew, K. S., & Flanagan, D. (1998). *The intelligence test desk reference*. Allyn and Bacon.

Mitchell, M. (2021). Abstraction and analogy-making in artificial intelligence. *Annals of the New York Academy of Sciences, 1505,* 79-101. http://dx.doi.org/10.1111/nyas.14619

Muñiz, J. & Fernández-Hermida, J. R. (2010). La opinión de los psicólogos españoles sobre el uso de los test [*The opinion of spanish psychologists on the use of tests*]. *Papeles del Psicólogo, 31,* 108-121.

O'Boyle, M. W., Cunnington, R., Silk, T. J., Vaughan, D., Jackson, G., Syngeniotis, A. & Egan, G. F. (2005). Mathematically gifted male adolescents activate a unique brain network during mental rotation. *Brain Research: Cognitive Brain Research, 25,* 583–587. http://dx.doi.org/10.1016/j.cogbrainres.2005.08.004

Panizzon, M. S., Fennema-Notestine, C., Eyler, L. T., Jernigan, T. L., Prom-Wormley, E., Neale, M., ... & Kremen, W. S. (2009). Distinct genetic influences on cortical surface area and cortical thickness. *Cerebral cortex*, *19*, 2728-2735. https://doi.org/10.1093/cercor/bhp026

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

Prabhakaran, V., Smith, J. A. L., Desmond, J. E., Glover, G. H., & Gabrieli, J. D. E. (1997). Neural Substrates of Fluid Reasoning: An fMRI Study of Neocortical Activation during Performance of the Raven's Progressive Matrices Test. *Cognitive Psychology, 33***,** 43–63.

Quintana, J., & Fuster, J. M. (1999) From perception to action: temporal integrative functions of prefrontal and parietal neurons. *Cerebral Cortex, 9,* 213–221. http://dx.doi.org/10.1093/cercor/9.3.213

Ramnani, N., & Owen, A. M. (2004). Anterior prefrontal cortex: Insights into function from anatomy and neuroimaging. *Nature Neuroscience, 5,* 184−194. http://dx.doi.org/10.1038/nrn1343

Ramsden, S., Richardson, F.M., Josse, G., Thomas, M. S. C., Ellis, C., Shakeshaft, C., Seghier, M. L., & Price, C. J. (2011). Verbal and nonverbal intelligence changes in the teenage brain. *Nature, 479,* 113 – 116. http://dx.doi.org/10.1038/nature10514

Raven, J. (1939). *Progressive Matrices: A perceptual test of intelligence*. H. K. Lewis.

Raven, J. (1962). *Advanced progressive matrices*. H. K. Lewis.

Raven, J., & Raven, J. (2008) *Uses and abuses of intelligence: Studied advancing Spearman and Raven's quest for non-arbitrary metrics*. New York: Royal Fireworks Press.

Román, F. J., Abad, F. J., Escorial, S., Burgaleta, M., Martínez, K., Álvarez-Linera, J., ... & Colom, R. (2014). Reversed hierarchy in the brain for general and specific cognitive abilities: A morphometric analysis. *Human Brain Mapping, 35,* 3805-3818. https://doi.org/10.1002/hbm.22438

Rowe, J. B., & Passingham, R. E. (2001). Working Memory for Location and Time: Activity in Prefrontal Area 46 Relates to Selection Rather than Maintenance in Memory. *Neuroimage*, *14*, 77-86. http://dx.doi.org/10.1006/nimg.2001.0784

Salgado-Pineda, P., Baeza, I., Perez-Gomez, M., Vendrell, P., Junque, C., Bargallo, N. & Bernardo, M. (2003). Sustained attention impairment correlates to gray matter decreases in first episode neuroleptic-naïve schizophrenic patients. *NeuroImage 19,* 365–375. http://dx.doi.org/10.1016/S1053-8119(03)00094-6

Shaw, P., Greenstein, D., Lerch, J., Clasen, L., Lenroot, R., Gogtay, N., Evans, A., Rapoport, J. & Giedd, J. (2006). Intellectual ability and cortical development in children and adolescents. *Nature, 440,* 676–79. http://dx.doi.org/10.1038/nature04513

Spearman, C. E. (1938). Measurement of intelligence. *Scientia Milano, 64,* 75–82.

Spearman, C., & Wynn-Jones, L. (1951). *Human ability*. London: Macmillan.

Sporns, O., Honey, C. J., & Kötter, R. (2007). Identification and Classification of Hubs in Brain Networks. *PLoS ONE 2(10)*: e1049. http://dx.doi.org/10.1371/journal.pone.0001049

Talairach, J., & Tournoux, P. (1988). *Co-planar stereotaxic atlas of the human brain: A 3-dimmensional proportional system, an approach to cerebral imaging*. New York: Thieme Medical Publishers.

Thiele, J. A., Faskowitz, J., Sporns, O., Chuderski, A., Jung, R., & Hilger, K. (2025). Decoding the Human Brain during Intelligence Testing. *bioRxiv*, 2025-2004. https://doi.org/10.1101/2025.04.01.646660

Thurstone, L. L., & Thurstone, Th. G. (1968). *Aptitudes mentales primarias (PMA) manual.* [Examiner manual for the SRA primary mental abilities test]. TEA Ediciones.

Torres, E. & Cuesta, M. (1992). Una revisión de las propiedades psicométricas del Test de matrices progresivas de Raven (Escala superior). [A Review of the Psychometric Properties of Raven's Progressive Matrices Test (Advanced Level)] *Psicothema, 4,* 261-267.

Vuoksimaa, E., Panizzon, M. S., Chen, C. H., Fiecas, M., Eyler, L. T., Fennema-Notestine, C., ... & Kremen, W. S. (2016). Is bigger always better? The importance of cortical configuration with respect to cognitive ability. *Neuroimage, 129*, 356-366. http://doi.org/10.1016/j.neuroimage.2016.01.049

Wager, T. D., & Smith, E. E. (2003). Neuroimaging studies of working memory: A meta-analysis. *Cognitive, Affective, and Behavioral Neuroscience, 3,* 255−274.

Webb, T., Holyoak, K.J. & Lu, H. (2023). Emergent analogical reasoning in large language models. *Nature Human Behavior, 7,* 1526–1541. https://doi.org/10.1038/s41562-023-01659-w

Wicherts, J. M., Dolan, C.V., Carlson, J.S., van der Maas, Han L. J. (2010). Raven's test performance of sub-Saharan Africans: Average performance, psychometric properties, and the Flynn Effect. *Learning and Individual Differences, 20,* 135-151. http://dx.doi.org/10.1016/j.lindif.2009.12.001

Wilhelm, O., Hildebrandt, A., & Oberauer, K. (2013). What is working memory capacity, and how can we measure it?. *Frontiers in psychology*, *4*, 433. https://doi.org/10.3389/fpsyg.2013.00433

Winfred, A. & Day, D. V. (1994). Development of a short form for the Raven Advanced Progressive Matrices Test. *Educational and Psychological Measurement, 54,* 394-403. http://dx.doi.org/10.1177/0013164494054002013

Yela, M. (1969). *Rotación de Figuras Macizas [Rotation of solid figures]*. TEA.

Zhao, S., You, H., Zhang, R. Y., Si, B., Zhen, Z., Wan, X., & Wang, D. H. (2023). An interpretable neuro-symbolic model for raven's progressive matrices reasoning. *Cognitive Computation*, *15*, 1703-1724. https://doi.org/10.1007/s12559-023-10154-3

Zilles, K., & Amunts, K. (2010). Centenary of Brodmann's map — conception and fate. *Nature Review Neuroscience, 11,* 139-145. http://dx.doi.org/10.1038/nrn2776

Zurrin, R., Wong, S. T. S., Roes, M. M., Percival, C. M., Chinchani, A., Arreaza, L., ... & Woodward, T. S. (2024). Functional brain networks involved in the Raven's standard progressive matrices task and their relation to theories of fluid intelligence. *Intelligence, 103,* 101807. https://doi.org/10.1016/j.intell.2024.101807

Appendix

Positive correlations (p < .001) between regional gray matter and measures of intelligence.

Intelligence measure		Brain Regions	x,y,z coordinates	cluster size
g	*Right Frontal*			
		BA 10 (Middle Frontal Gyrus)	34, 42, 6	47
		BA 8 (Middle Frontal Gyrus)	32, 12, 44	19
		BA 4 (Precentral Gyrus)	20, -38, 76	19
Gf	*Right Frontal*			
		BA 8 (Middle Frontal Gyrus)	32, 12, 44	17
Gc	*Right Frontal*			
		BA 8 (Middle Frontal Gyrus)	32, 14, 44	27
		BA 10 (Middle Frontal Gyrus)	34, 42, 6	27
Gv	*Right Frontal*			
		BA 10 (Middle Frontal Gyrus)	34, 42, 6	36
	Right Temporal			
		BA 37 (Fusiform Gyrus)	44, -38, -14	15
	Right Sub-lobar			
		Caudate (Body)	12, 16, 8	88
	Left Sub-lobar			
		Caudate (Head)	-14, 20, 6	88
RAPM	*Right Frontal*			
		BA 8 (Middle Frontal Gyrus)	32, 12, 44	33
		BA 9 (Middle Frontal Gyrus)	32, 34, 28	11
	Left Sub-lobar			
		Caudate (Body)	-12, 4, 16	39
PMA-R	*Left Sub-lobar*			
		BA 13 (Insula)	-38, -42, 16	72
DAT-AR(*)				
PMA-V	*Right Frontal*			
		BA 46 (Inferior Frontal Gyrus)	33, 32, 16	18
		BA 11 (Middle Frontal Gyrus)	30, 42, 0	12
	Left Frontal			
		BA 47 (Inferior Frontal Gyrus)	-20, 34, -6	13
DAT-VR				
	Right Sub-lobar			
		Putamen (Lentiform Nucleus)	32, -22, 2	241
DAT-NR	*Right Frontal*			
		BA 10 (Middle Frontal Gyrus)	34, 42, 10	53
		BA 6 (Paracentral Lobule)	10, -32, 56	53
		BA 8 (Middle Frontal Gyrus)	54, 14, 46	215
	Right Parietal			
		BA 5 (Postcentral Gyrus)	24, -46, 68	124
		BA 7 (Precuneus)	14, -74, 50	64
		BA 40 (Inferior Parietal Lobule)	50, -30, 30	12
	Right Temporal			
		BA 39 (Middle Temporal Gyrus)	54, -66, 26	85
	Right Occipital			
		BA 19 (Supperior Occipital Gyrus)	54, -66, 26	659
	Left Frontal			
		BA 6 (Medial Frontal Gyrus)	-8, -26, 54	63

Intelligence measure		Brain Regions	x,y,z coordinates	cluster size
		BA 6 (Medial Frontal Gyrus)	0, -6, 62	60
	Left Occipital			
		BA 19 (Middle Occipital Gyrus)	-54,-70,-12	60
		BA 19 (Superior Occipital Gyrus)	-38, -88, 32	41
Sol. Figures(*)				
DAT-SR	*Left Temporal*			
		BA 21 (Middle Temporal Gyrus)	-50, -32, 6	15
	Left Limbic Lobe			
		BA 23 (Cingulate Gyrus)	-6, -14, 30	23
PMA-S	*Right Frontal*			
		BA 25 (Subcallosal Gyrus)	6, 6, 20	118
	Right Sub-lobar			
		Caudate (Head)	10, 20, 2	15
	Left Cerebellum			
		Culmen	-16, -24, -28	52

Note. (*) Not found clusters bigger than 10mm. Brain regions (approximate Brodmann areas, BAs) are estimated from Talairach and Tournoux (1988) atlas. Coordinates refer to maximum voxel of identified clusters. Cluster size indicates number of voxels with a significant correlation with intelligence, excluding cluster sizes < 10 mm. Overlap between BA 8 and measures of intelligence (red). Overlap between BA 10 and measures of intelligence (blue).