



Published in final edited form as:

*Intelligence*. 2013 March ; 41(2): 129–140. doi:10.1016/j.intell.2013.01.002.

## Hippocampal structure and human cognition: key role of spatial processing and evidence supporting the efficiency hypothesis in females

Roberto Colom<sup>1</sup>, Jason L. Stein<sup>2</sup>, Priya Rajagopalan<sup>2</sup>, Kenia Martínez<sup>1</sup>, David Hermel<sup>2</sup>, Yalin Wang<sup>2</sup>, Juan Álvarez-Linera<sup>3</sup>, Miguel Burgaleta<sup>1</sup>, M<sup>a</sup>Ángeles Quiroga<sup>4</sup>, Pei Chun Shih<sup>1</sup>, and Paul M. Thompson<sup>2</sup>

<sup>1</sup>Facultad de Psicología, Universidad Autónoma de Madrid, 28049, Madrid (SPAIN) and Fundación CIEN-Fundación Reina Sofía, C/ Valderrebollo, 5, 28031, Madrid (SPAIN)

<sup>2</sup>Laboratory of Neuroimaging (LONI), UCLA School of Medicine, 635 Charles E. Young Drive South, Suite 225, Los Angeles, CA 90095-7334 (USA)

<sup>3</sup>Hospital Ruber Internacional-Fundación CIEN-Fundación Reina Sofía, C/ Valderrebollo, 5, 28031, Madrid (SPAIN)

<sup>4</sup>Facultad de Psicología, Universidad Complutense de Madrid, Campus de Somosaguas, 28223, Pozuelo de Alarcón, Madrid (SPAIN)

### Abstract

Here we apply a method for automated segmentation of the hippocampus in 3D high-resolution structural brain MRI scans. One hundred and four healthy young adults completed twenty one tasks measuring abstract, verbal, and spatial intelligence, along with working memory, executive control, attention, and processing speed. After permutation tests corrected for multiple comparisons across vertices ( $p < .05$ ) significant relationships were found for spatial intelligence, spatial working memory, and spatial executive control. Interactions with sex revealed significant relationships with the general factor of intelligence ( $g$ ), along with abstract and spatial intelligence. These correlations were mainly positive for males but negative for females, which might support the efficiency hypothesis in women. Verbal intelligence, attention, and processing speed were not related to hippocampal structural differences.

### Keywords

Hippocampus; Intelligence; Working memory; Executive control; Attention; Processing speed; Sex differences

## INTRODUCTION

The hippocampus is a small brain region, involved in cognition and emotion, capturing a lot of attention from the scientific community (Pessoa, 2008). Hippocampus structure and function has been related to spatial processing (Frings et al., 2006, Maguire et al., 2000), working memory (Axmacher et al., 2010), episodic memory, and the construction of mental images (Bird, & Burgess, 2008, Squire et al., 2010) among other cognitive factors. However, the available evidence is mainly based on narrow cognitive measures, small and heterogeneous samples, or broad anatomical hippocampal estimates. Obtaining a comprehensive view regarding the relationships between hippocampus structure and human cognition in general requires (a) the analysis of large and well defined samples, (b) diverse measures of the same cognitive factor, and (c) fine-grained hippocampal maps. Neglecting these conditions opens the door to incomplete and inconsistent findings.

To further complicate things, this brain structure is highly complex. Thus, for instance, with respect to spatial processing, Maguire et al. (2000) found no differences in overall hippocampal volume between taxi drivers and controls. However, drivers showed larger posterior hippocampi, whereas controls had larger anterior hippocampal volumes. Also, it has been shown that not all working memory components are related to the hippocampus: it is not relevant for working memory binding processes (Baddeley, Allen, & Vargha-Khadem, 2010) and it is more active for associating items from different cortical regions than from the same cortical region (Piekema et al., 2009). The hippocampus contributes to a temporo-occipital network for relational working memory maintenance, but to a fronto-parietal network for non-relational maintenance (Cashdollar et al., 2009, Nee, & Jonides, 2011).

Therefore, it can be stated that the hippocampus is not a homogeneous structure (Pantel et al., 2000). The anterior region is mainly, but not exclusively, related to emotion—it is connected with the prefrontal cortex, the amygdala, and the hypothalamic-pituitary-adrenal axis—whereas the posterior region is essentially implicated in cognition (Bannerman et al., 2004). Also, regional structural changes across development (age changes) have been documented: the anterior and posterior regions are reduced and increased over time respectively. Volume reduction at the posterior regions appears to be greater for females, while volume reduction at the anterior region is larger in males (Gogtay et al., 2006). Cahill (2006) underscored the need for challenging the presumption that sex is not important in neuroscience research, specifically discussing data related to the hippocampus. He underscores that male and female hippocampi are significantly different in anatomical structure, but "*whether and how sex influences hippocampal function has not been yet systematically examined*" (p. 3).

As noted above, most available studies consider isolated cognitive tests or tasks, as well as arguable defined general estimates, for relating cognitive performance differences to hippocampal structure and function. Sometimes a significant relationship is found for a given cognitive function, whereas other times researchers fail to obtain significant results. Contradictory or inconsistent findings may be attributed to several factors, such as the specific administered tasks for estimating the construct of interest or sample's characteristics (size, sex, age, etc.). Overcoming this situation involves the analysis of large and well

defined samples tapping the constructs of interest as representatively as possible administering several tasks (Barbey et al., 2012; Colom et al., 2009, 2010; Colom, & Thompson, 2011; Gläscher et al., 2010, Haier et al., 2009).

For this purpose here we analyze the relationships between regional hippocampal structural differences and a wide set of cognitive factors, namely abstract-fluid, verbal-crystallized, and spatial intelligence, along with working memory, executive control, attention, and processing speed. Fluid-abstract intelligence (Gf) assesses the complexity level that subjects can handle in situations for which previous knowledge is not relevant, whereas crystallized-verbal intelligence (Gc) relies in the ability to cope with academic types of skills and knowledge, such reading or arithmetic (Cattell, 1971). Spatial intelligence (Gv) implicates the construction, temporary retention, and manipulation of mental images (Lohman, 2000). The general factor of intelligence (*g*) is a source of variance common to a variety of abstract, verbal, and non-verbal abilities (Gläscher et al., 2010, Karama et al., 2011). Working memory can be defined as the ability for the simultaneous storage and processing of varied amounts of information (Colom et al., 2006). Executive control implicates the ability for regulating mental processes. Inhibition, shifting or updating are key components of this type of control (Friedman et al., 2006). Attention is a broad cognitive function for focusing available mental resources (Baddeley, 2002). Here we consider the control of automatic responses (inhibition). Finally, processing speed is usually measured by reaction time tasks (Sheppard & Vernon, 2008), so simple verification tasks are administered in the present study. Figure 1 depicts examples of tests and tasks measuring some of these cognitive factors (detailed descriptions can be found below). While there are published studies considering some of these functions in isolation, to our knowledge there is none applying automated mapping of hippocampal structure to a large and well defined sample of participants assessing the constructs of interest by diverse cognitive tests and tasks.

## METHOD

### Participants and neuroimaging data

The sample comprised 104 young adults (59 females and 45 males) with a mean age of 19.9 (SD = 1.6, age range = 18–27)\*. They were selected to avoid medical or psychiatric illness, including a history of head injury and substance abuse. This was carefully checked by a standardized questionnaire regularly administered where the scanning took place (Fundación Reina Sofía, Madrid). Written informed consent was obtained following the Helsinki guidelines (World Medical Association, 2008).

MRIs were obtained with a 3T scanner\*\*. The scans were lineally registered to the ICBM template with a 9 parameter transformation (3 translations, 3 rotations, 3 scales) using the

\*This sample was previously analyzed for achieving quite different research goals. Colom et al. (2009) applied a VBM protocol for analyzing gray matter correlates of fluid, crystallized, and spatial intelligence factors (excluding the cerebellum) in 100 participants from this sample. Burgaleta et al. (2012) analyzed sex differences in this same one hundred participants exploring total, gray, and white matter volumes with respect to abstract, verbal, and spatial intelligence subtests. Bruner et al. (2010, 2011, 2012), and Martín-Loeches et al. (in press) based their research in 2D midsagittal brain shape variations.

\*\* 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). Sag acquisition .8 mm thickness, full brain coverage (220 slices), matrix 266×266 FOV 24 (isotropic voxels .7 cm<sup>3</sup>).

Mintracc algorithm. Aligned images were resampled in an isotropic space with a voxel size of  $1 \text{ mm}^3$ .

Once the 3D registered volumes were available, twenty representative cases in terms of sex and variability in cognitive indices were chosen for manual tracing. These cases were manually traced by two researchers that worked independently with the forty images until they achieved appropriate reliability indices after the delineated labels: sensitivity = .84, specificity = 1, overlay = .78. Next the remaining eighty four (testing) brains were submitted to a LONI Pipeline for their automatic segmentation (Morra et al., 2008).

This Pipeline comprises several steps and produces outputs for the left and right hippocampus. Afterwards a detailed quality control (QC) was performed for checking these outputs, as detailed below. A second Pipeline was then applied for obtaining hippocampal surface indices. Computed regional volumetric differences for the left and right hippocampus were related to the cognitive scores of interest. Probability maps were obtained after permutation tests corrected for multiple comparisons ( $p < .05$ ). Specific details regarding manual delineation of the hippocampus, the automatic recognition algorithm, and the production of statistical maps can be found below.

### Psychological measures

Twenty one cognitive tests and tasks were administered for measuring the psychological constructs of interests defined above. Importantly, all these constructs were estimated by three different measures for obtaining theoretically representative scores.

Abstract-fluid intelligence (Gf) was measured by the Advanced Progressive Matrices Test (APM) (Raven, Raven & Court, 2004), the abstract reasoning subtests from the Differential Aptitude Test (DAT-AR) battery (Bennett, Seashore & Wesman, 2005), and the inductive reasoning subtests from the Primary Mental Abilities (PMA-R) battery (Thurstone, 1938). The APM comprises a matrix figure with three rows and three columns. Among eight possible alternatives the one completing the  $3 \times 3$  matrix figure must be chosen. The screening version comprising odd items only was administered (max. score = 18). DAT-AR is a series test based on abstract figures. Successive figures follow a given rule, so the one continuing the series must be chosen from several alternatives. The screening version comprising odd items only was administered (max. score = 20). PMA-R comprises letters' series items. The rule (or rules) underlying a given sequence must be extracted for selecting the correct alternative (max. score = 30).

Verbal-crystallized intelligence (Gc) was measured by the verbal and numerical reasoning subtests from the DAT (VR and NR), along with the vocabulary subtests from the PMA (V). DAT-VR is based on sentences stated like an analogy. The first and last words from the sentence are missing, and a pair of words completing the sentence must be selected. The screening version comprising odd items only was administered (max. score = 20). PMA-V is a synonym test based on the meaning of words that must be evaluated against a given model word (max. score = 50). DAT-NR consists of quantitative reasoning problems. The screening version comprising odd items only was administered (max. score = 20).

Spatial intelligence (Gv) was measured by the spatial relations subtests from the DAT (SR), the spatial rotation subtests from the PMA (S), and the rotation of solid figures test (Yela, 1969). Items from the rotation of solid figures test are based on a 3D model figure and several 3D rotated alternatives (max. score = 21). PMA-S includes a model figure and six alternatives, some of whom are simply rotated versions of the model figure, whereas the remaining figures are mirror imaged. Only rotated figures are to be selected and there could be several correct options for each item. The score is the total number of correct responses (appropriately selected figures – simply rotated) minus the total number of incorrect responses (inappropriately selected figures – mirror imaged) (max. score = 54). DAT-SR is a mental folding test. One unfolded figure is shown at the left, whereas figures at the right depict folded versions. The folded figure matching the unfolded figure at the left must be chosen. The screening version comprising odd items only was administered (max. score = 25).

Working memory (WM) was measured by the reading span, computation span, and dot matrix tasks (Colom et al., 2010). In the reading span task participants verify if a set of sentences sequentially displayed make or make no sense. Each display includes a sentence and a to-be remembered capital letter. Sentences are 10 – 15 words long. At the end of a given set, participants recall, in their correct serial order, each letter from the set. Set sizes range from 3 to 6 sentence/letter pairs per trial, for a total of 12 trials (4 levels  $\times$  3 trial = 12 trials total). The computation span task includes a verification task and a recall task. 6 seconds are allowed to see the math equation without a time limit for verifying its accuracy. The displayed solution, irrespective of its accuracy, must be serially remembered at the end of a given set. Each math equation includes two operations using digits from 1 to 10. The solutions are always single-digit numbers. Trials range from three to seven equation/solutions (5 levels  $\times$  3 trials each = 15 trials total). In the dot matrix task, a matrix equation must be verified and a dot location displayed in a five  $\times$  five grid must be retained. The matrix equation is presented during a maximum of 4.5 seconds for adding or subtracting simple line drawings. Once the response is given, the grid comprising the to-be remembered dot is displayed for 1.5 s. After a given set of equation-grid pairs, the grid spaces that contained dots must be recalled clicking with the mouse on an empty grid. Trials increase in size from two to five equations and dots (4 levels  $\times$  3 trials = 12 trials total). The score for these three WM tasks is the number of hits in the verification and recalling tasks.

Executive control (updating, UPD) was measured by the 2 back, keep track, and letter memory tasks (Colom et al., 2008). In the 2-back task upper and lower case letters are presented in one of eight equidistant spatial locations around the center of a computer screen. Stimuli are presented for 200 ms and 1300 ms are given for responding. There are 75 experimental stimuli of which 25 are match stimuli. Participants press the space bar of the keyboard to make a match response (a letter presented in the same spatial location 2 positions back in the sequence). The score is the number of hits. In the keep track task participants see several categories at the bottom of the computer screen. Fifteen items, including two or three exemplars from each of the six possible categories (Odd, even, vowel, consonant, lowercase pairs of letters, and uppercase pairs of letters) are then presented serially and in random order for 1500 ms each, with the target categories remaining at the bottom of the screen. The last item presented for each target category must be retained.

Therefore, items must be monitored for updating memory representations for the appropriate categories. After the practice trials, participants complete three trials with four target categories, and three trials with five target categories, recalling a maximum of 27 items. In the letter memory task several letters are presented serially for 1000 ms per letter. The last four letters presented in the list must be remembered. Instructions require and emphasize rehearsing the last four letters by mentally adding the most recent letter and deleting the fifth letter back. Participants perform three practice trials, and there are six experimental trials of varying length (15, 17, 19, 21, 23, 25) randomly presented, for a total of 24 letters recalled.

Attention (ATT) was measured by verbal and numerical versions of the flanker task, along with a spatial variant of the Simon task (Colom et al., 2010). The verbal and quantitative tasks require deciding, as fast as possible, if the letter/digit presented in the center of a set of three letters/digits is vowel/odd or consonant/even. The target (e.g. vowel/odd) can be surrounded by compatible (e.g. vowel/odd) or incompatible (e.g. consonant/even) letters/digits. The spatial task requires deciding if an arrow (horizontally depicted) points to the left or to the right of a fixation point. The target arrow pointing to a given direction (e.g. to the left) can be presented at the left (e.g. compatible) or at the right (e.g. incompatible) of the fixation point. There are a total of 32 practice trials and 80 experimental trials. Half of the trials are compatible and they are randomly presented across the entire session. The mean reaction time for the incompatible trials is the dependent measure.

Finally, processing speed (PS) was measured by simple recognition verbal, numerical, and spatial tasks (Colom et al., 2008). In these speed tasks one or two items (letter, digit, or arrows) are sequentially displayed for 650 ms. each. Those items define a memory set that can comprise (a) uppercase and lowercase letters, (b) digits, or (c) arrows with different shapes. After the last displayed item, a fixation point appears for 500 ms. Finally, the probe item appears in order to decide, as quickly and accurately as possible (a) if it has the same *meaning* as one of the letters, (b) if the number can be divided by one of the digits, or (c) if it has the same orientation of one of the arrows, presented within the memory set. Experimental trials range from one to two items (2 levels  $\times$  30 trials each = 60 trials total). For example [Verbal speed]: [Memory Set] [First screen] A >>> [Second screen] a >>> [Answer screen, Same meaning?] 'Yes' 'No'. The score is the mean RT for the correct answers.

All these tests and tasks were always administered in the same order. Intelligence tests were administered in sessions 1 and 2, whereas cognitive tasks were administered in sessions 3 and 4. Session 1 was devoted to the RAPM (20 min), inductive reasoning (PMA-R, 6 min), vocabulary (PMA-V, 4 min), and abstract reasoning (DAT-AR, 10 min), while session 2 was comprised by verbal reasoning (DAT-VR, 10 min), Rotation of solid figures (5 min), numerical reasoning (DAT-NR, 10 min), mental rotation (PMA-S, 5 min), and spatial relations (DAT-SR, 10 min). Session 3 included the working memory and processing speed tasks, whereas session 4 was comprised by the executive and attention tasks.

### Manual delineation of the Hippocampus

The training brains set comprised twenty subjects (half female) chosen for manual delineation of the left and right hippocampus (forty images). Manual delineation of the left

and right structures was performed on the template registered images using BrainSuite 2.0 (Shattuck & Leahy, 2002) by two researchers working independently. Reliability of the manual traces was assessed by sensitivity (the likelihood that a voxel defined by researcher A as being within the hippocampus has also been defined by researcher B as being within the hippocampus), specificity (the likelihood that a voxel defined by one researcher as not being within the hippocampus has also not been defined by the other researcher as being within the hippocampus), and overlap (the percentage of voxels classified as hippocampus by both researchers divided by the voxels classified as hippocampus by either researcher).

Hippocampal delineation was conducted according to guidelines provided by Duvernoy (1998) and Pantel et al. (2000). The hippocampus includes the cornu ammonis subfields (CA1-4), dentate gyrus, and subiculum. Manual tracing was started on the sagittal plane, moving from lateral to medial. The first slice was chosen where the gray matter of the hippocampus opened in a pocket of the lateral ventricles. Tracing continued medially, excluding the ventricles, fimbria, and surrounding white matter. In anterior portions, the hippocampus was distinguished from the amygdala through small contrast change of the alveus surrounding the hippocampal head. The alveus was excluded whenever possible moving between slices and between sagittal and coronal views along with references to neuroanatomical delineations (Duvernoy, 1998).

To reinforce the smoothness of the brain in the manual delineation, structures were reviewed and edited in the coronal plane moving from posterior to inferior sections. Jumps in boundaries between slices could be visualized as jagged borders in a perpendicular plane. Afterwards, the manual tracing went back to the sagittal plane for re-editing and ensuring the pursued smoothness.

### Automatic Recognition

The 20 left and 20 right training brains were separately used to train an automated recognition algorithm to delineate the structure in the remaining 84 brains. The so-called auto-context model, based on a modified AdaBoost algorithm (Morra et al., 2008) was run. This method takes in the binary training scans of one researcher, template registered structural MRI images, and tissue segmented images of the 20 manually delineated brains. It then extracts features from the MRI and tissue segmented images based on image intensity, position, image curvatures, image gradients, and Haar filters. Combinations of features are used to predict the probability of each voxel belonging to a defined structure in each of the 84 testing brains. Voxels with greater than 0.5 probability are identified as being within the structure of interest. Each automated structural definition was explicitly checked, so stray voxels and false classifications (i.e. hippocampus extending into parahippocampal gyrus, contralateral putamen being identified, etc.) were manually corrected. Small islands of voxels within slices which were disconnected from other voxels of the structure were also removed.

### Statistical maps

Statistical maps are generated to reveal local differences in radial hippocampal distance and individual differences in these local surface parameters. At each hippocampal point multiple

regression evaluates whether individual differences at that point are related to the covariate of interest. The  $p$  value describing the significance of this association controlling for nuisance covariates (sex, age, and handedness for the complete group) was plotted onto the surface at each point on the hippocampus using a colour code to produce a statistical map. Overall  $p$  values were assigned to the maps using a permutation testing approach which measures the distribution of a given feature in statistical maps (the area with statistics above a predetermined threshold) that would be observed by accident if subjects were randomly assigned to the covariates of interest. This distribution is used to compare the features that occurred empirically with those that occurred by accident in the random assignments. A ratio is computed describing what fraction of the time an effect of similar or greater magnitude to the real effect occurs in the random assignments. This is the chance of the observed pattern occurring by accident. This fraction provides an overall significance value ( $p$  value) for the whole map. The proportion of the surface where individual differences are related to the covariates of interest at a primary threshold of  $p < 0.05$  was assessed.

One sided tests were used, so positive  $p$ -maps depict a direct relationship –radial distance increases with score increase in the covariate of interest—whereas negative  $p$ -maps indicate an inverse relationship –radial distance decreases with increase in the covariate of interest.

Two further contrasts were applied. First, the interaction among sex, cognitive scores, and hippocampal structural differences (controlling for age and handedness). Afterwards, beta maps were computed for significant findings derived from the above interaction. This was intended for interpreting areas of significant interaction. The main question to answer here is how the hippocampus is associated with cognitive scores between males and females.

## RESULTS

Results for the psychological constructs/measures separated by sex are shown in Table 1. Effect sizes ( $d$ ) are also reported in the last column. It can be seen that females outperform males in PMA-R (inductive reasoning), keep track (executive updating), and numerical attention, whereas males outperform females in the rotation of solid figures test (spatial intelligence) only. Anyways, most of the differences are not significant, and certainly they are not in the most general estimates of cognition (eight first rows).

Simultaneous relationships among the assessed psychological constructs were tested through structural equation modelling (SEM) --Figure 2. Fit indices for this model were reasonable: CMIN = 242.4, DF = 168, CMIN/DF = 1.4, RMSEA = .065. There are several noteworthy results. First, working memory (WM) and updating (UP) show a perfect correlation. Second, fluid intelligence (Gf) and updating show a very high correlation (.93). Third, processing speed is related to crystallized and spatial, but not to fluid intelligence. Fourth, attention is only related to processing speed and updating. Finally, most of the correlations are statistically significant (although quite different in magnitude) but there are also non-significant relationships (those not shown in Figure 2).

With respect to imaging analyses, here we report statistical maps build from significant relations between hippocampal regional structural differences and the psychological



constructs/measures of interest after the computation of permutation tests corrected for multiple comparisons ( $p < .05$ ) as noted above.

First, the complete group ( $N = 104$ ) was considered (controlling for sex, age, and handedness). As depicted in Figure 3, significant relationships were mainly found in posterior areas for general spatial intelligence ( $G_v$ ), the spatial relations (SR) subtest from the Differential Aptitude Test (DAT) battery (DAT-SR), the spatial n back task measuring executive control (2 back), and the spatial working memory task (dot matrix).

Secondly, sex differences in hippocampal structure were inspected. Top panel of Figure 4 shows total, right, and left volumes for males and females (effect sizes,  $d$ , are also depicted), whereas the bottom panel depicts regional differences in the right and left hippocampus. Males do show larger hippocampi, especially on the left side ( $d = .56$ ). Regionally, sex differences are concentrated on the posterior regions of the right hippocampus, but they are more widespread on the left side.

Third, because of the observed sex differences, interactions among hippocampal structure, cognitive measures, and sex were computed (controlling for age and handedness). Results can be inspected in Figure 5.

Significant findings were observed for the general factor of intelligence ( $g$ ), abstract-fluid intelligence ( $G_f$ ), and the Abstract Reasoning subtest from the DAT battery (DAT-AR) on the left posterior hippocampus, as well as for the Advanced Progressive Matrices (APM) test and the rotation of solid figures test on the right anterior hippocampus. The right posterior hippocampus is also relevant for the rotation of solid figures test.

Beta maps, depicted on the right panel of Figure 5, show mainly positive relationships for males but negative for females. These beta maps allow interpreting areas of significant interaction regarding how the hippocampus is associated with cognitive scores between sexes. Thus, there is a positive association with hippocampal shape and cognitive performance in males --the larger the radial distance, the better the score--, whereas for females this association is negative --the shorter the radial distance, the better the performance.

Finally, verbal intelligence, attention, and processing speed scores were unrelated to hippocampal regional structural differences. Therefore, verbal-crystallized abilities, controlled inhibition processing, and mental speed processes, were not significantly associated to hippocampal structure in this sample.

## DISCUSSION

Here we report findings providing what might be considered an interesting general picture. Firstly, varied --but not general-- measures comprising spatial processing requirements are significantly related to hippocampal structural differences. Specifically, spatial intelligence, spatial executive control processes, and spatial working memory. It is noteworthy that abstract, verbal, and numerical processing were not significantly related to this brain structure for these contrasts controlling for sex, age, and handedness. The n back task

revealed significant associations with the left and right hippocampus, the dot matrix task was circumscribed to the left posterior hippocampus, and spatial ability was mainly focused on the right hippocampus.

Secondly, these findings change when analyzing the interaction with sex. Now the general factor of intelligence ( $g$ ) --comprised by variance common to abstract, verbal, and spatial intelligence-- as well as abstract-fluid intelligence ( $G_f$ ) and abstract reasoning (DAT-AR) are associated with the left posterior hippocampus only. However, the Advanced Progressive Matrices (APM) test and the rotation of solid figures test are related to the right hippocampus only.

Finally, the computed beta maps revealed that these significant associations were mainly positive for males but negative for females, suggesting higher efficiency for women (Haier et al., 1988, 2005, Neubauer & Fink, 2009 a). The efficiency hypothesis states that people showing better cognitive performance do so along with lower brain activation than people with worse cognitive scores. However, the evidence regarding sex is largely confusing, as noted by Neubauer and Fink (2009 b). Haier and Benbow (1995) failed to find positive evidence for the neural efficiency hypothesis, but they reported substantial sex differences. Neubauer et al. (2002) reported neural efficiency for males in spatial tasks, but for females in verbal tasks. Tang et al. (2010) found positive correlations between intelligence and integrity of interhemispheric white matter connections in females, but these correlations were negative in males. The findings reported here might help providing some light because, as discussed by Neubauer and Fink (2009 b), functional results may or may not fit structural ones. At this structural level, females might show greater efficiency requiring less neural material for achieving behavioural results on a par with males.

As underscored above, the hippocampus is a small brain structure, but it is also largely complex. Hippocampus atrophy is of the utmost interest in Alzheimer research, but it has been observed that, across time, MCI patients can improve their mental status, develop AD, or remain within this diagnostic category (Apostolova et al., 2006). Hippocampus volume is sensitive to aerobic exercise, but the positive impact is focused on the anterior region of the left and right sides (Erickson et al., 2011). Children raised in low income families do show smaller hippocampal volumes than children from medium and high income homes (Hanson et al., 2011). These results are just examples consistent with the complexities related to this brain region.

Hines (2011) noted that "*there are many reports of sex differences in the human brain and many of these are yet unreplicated (...) some findings of neural sex differences may prove to be spurious*" (p. 77). Haier et al. (2005) reported substantial sex differences in the structural correlates of IQ, but Schmidt et al. (2009) showed that males and females performance on the n back task is supported by essentially the same brain networks. These two examples nicely fit Hines's suspicion, but also our previous claim regarding the inconsistent findings that might be derived from the lack of analyses of a wide variety of cognitive measures with the same sample.

In conclusion, the present comprehensive analyses of the relationships between hippocampal structural differences and general human cognition suggests that the posterior sections of the hippocampus are relevant for several, but not general, spatial processing requirements. This statement applies to classic psychometric measures of spatial ability, but also to spatial working memory, and spatial executive control. However, spatial processing speed and spatial attention performance is not related to the studied hippocampal differences. This big picture changes when the interaction with sex is explicitly addressed. In such a case, the relevance of spatial processes is circumscribed to the rotation of solid figures test, and general (*g*) and abstract-fluid (*Gf*) intelligence become significant. Interestingly, for *g* and *Gf* significant findings were found on the left posterior hippocampus, while this same region is relevant for abstract reasoning (DAT-AR) also, but not for the advanced progressive matrices test (APM). The fact that both the APM and the rotation of solid figures test were related to the right hippocampus might be consistent with the widely recognised --but frequently neglected-- APM spatial requirements (Ackerman, Beier, & Boyle, 2002).

## Acknowledgments

This work was supported by grants SEJ-2006-07890, PR2008-0038, and PSI2010-20364. Software for automated segmentation was developed at LONI (Laboratory of Neuroimaging), School of Medicine, UCLA.

## REFERENCES

- Ackerman PL, Beier ME, Boyle MO. Individual differences in working memory within a nomological network of cognitive and perceptual speed abilities. *Journal of Experimental Psychology: General*. 2002; 131(4):567–589. [PubMed: 12500864]
- Apostolova LG, Dutton RA, Dinov ID, Hayashi KM, Toga AW, Cummings JL, Thompson PM. Conversion of Mild Cognitive Impairment to Alzheimer Disease Predicted by Hippocampal Atrophy Maps. *Archives of Neurology*. 2006; 63:693–699. [PubMed: 16682538]
- Axmacher N, Henseler MM, Jensen O, Weinreich I, Elger CE, Fell J. Cross-frequency coupling supports multi-item working memory in the human hippocampus. *PNAS*. 2010; 107(7):3228–3233. [PubMed: 20133762]
- Baddeley A. Is working memory still working? *European Psychologist*. 2002; 7:85–97.
- Baddeley A, Allen R, Vargha-Khadem F. Is the hippocampus necessary for visual and verbal binding in working memory? *Neuropsychologia*. 2010; 48:1089–1095. [PubMed: 20006631]
- Bannerman DM, Rawlins JN, McHugh SB, Deacon RM, Yee BK, Bast T, Zhang WN, Pothuizen HH, Feldon J. Regional dissociations within the hippocampus—Memory and anxiety. *Neuroscience Biobehavioural Review*. 2004; 28:273–283.
- 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. [PubMed: 22396393]
- Bennett, GK.; Seashore, HG.; Wesman, AG. *Differential Aptitude Test (DAT-5)*. Madrid, TEA, S.A.: 2005.
- Bird CM, Burgess N. The hippocampus and memory: Insights from spatial processing. *Nature Reviews, Neuroscience*. 2008 Mar.9:182–194.
- Bruner E, Cuétara JM, Colom R, Martín-Loeches M. Gender-based differences in the shape of the human corpus callosum are associated with allometric variations. *Journal of Anatomy*. 2012; 220:417–421. [PubMed: 22296183]
- Bruner E, Martín-Loeches M, Burgaleta M, Colom R. Midsagittal brain shape correlation with intelligence and cognitive performance. *Intelligence*. 2011; 39:141–147.
- Bruner E, Martín-Loeches M, Colom R. Midsagittal brain shape variation: patterns, allometry, and integration. *Journal of Anatomy*. 2010; 216:589–599. [PubMed: 20345859]

- Burgaleta M, Head K, Álvarez-Linera J, Martínez K, Escorial S, Haier R, Colom R. Sex Differences in Brain Volume Related to Specific Skills, not to General Intelligence. *Intelligence*. 2012; 40:60–68.
- Cahill, L. *Nature Reviews Neuroscience*. AOP: 2006. Why sex matters for neuroscience. published online 10 May 2006;
- Cashdollar N, Malecki U, Rugg-Gunn FJ, Duncan JS, Lavie N, Duzel E. Hippocampus-dependent and -independent theta-networks of active maintenance. *PNAS*. 2009; 106(48):20493–20498. [PubMed: 19918077]
- Cattell, RB. *Abilities: Their Structure, Growth, and Action*. Boston, Houghton-Mifflin: 1971.
- Colom R, Abad FJ, Quiroga M<sup>a</sup>A, Shih PC, Flores-Mendoza C. Working memory and intelligence are highly related constructs, but why? *Intelligence*. 2008; 36:584–606.
- Colom R, Haier RJ, Head K, Álvarez-Linera J, Quiroga M<sup>a</sup>A, Shih PC, Jung RE. Gray matter correlates of fluid, crystallized, and spatial intelligence: Testing the P-FIT model. *Intelligence*. 2009; 37:124–135.
- Colom R, Karama S, Jung RE, Haier RJ. Human intelligence and brain networks. *Dialogues in Clinical Neuroscience*. 2010; 12:489–501. [PubMed: 21319494]
- Colom R, Quiroga M<sup>a</sup>A, Shih PC, Martínez K, Burgaleta M, Martínez-Molina A, Román FJ, Requena L, Ramírez I. Improvement in working memory is not related to increased intelligence scores. *Intelligence*. 2010; 38:497–505.
- Colom R, Rebollo I, Abad F, Shih PC. Simple span tasks, complex span tasks, and cognitive abilities: A re-analysis of key studies. *Memory & Cognition*. 2006; 34:158–171. [PubMed: 16686115]
- Colom, R.; Thompson, PM. Understanding Human Intelligence by Imaging the Brain. In: Chamorro-Premuzic, T.; Furnham, A.; von Stumm, S., editors. *The Wiley-Blackwell Handbook of Individual Differences*. London: Wiley-Blackwell; 2011. p. 330-352.
- Duvernoy, HM. *The Human Hippocampus*. Heidelberg: Springer-Verlag; 1998.
- Erickson KI, Voss MW, Prakash RS, Basak C, Szabo A, Chaddock L, Kim JS, Heo S, Alves H, White SM, Wojcicki TR, Mailey E, Vieira VJ, Martin SA, Pence BD, Woods JA, McAuley E, Kramer AF. Exercise training increases size of hippocampus and improves memory. *PNAS*. 2011 Feb 15; 108(7):3017–3022. [PubMed: 21282661]
- Friedman NP, Miyake A, Corley RP, Young SE, DeFries JC, Hewitt JK. Not all executive functions are related to intelligence. *Psychological Science*. 2006; 17:172–179. [PubMed: 16466426]
- Frings L, Wagner K, Unterrainer J, Spreer J, Halsband U, Schulze-Bonhage A. Gender-related differences in lateralization of hippocampal activation and cognitive strategy. *Neuroreport*. 2006; 17(4):417–421. [PubMed: 16514369]
- Gläscher J, Rudrauf D, Colom R, Paul LK, Tranel D, Damasio H, Adolphs R. The distributed neural system for general intelligence revealed by lesion mapping. *PNAS*. 2010; 107(10):4705–4709. [PubMed: 20176936]
- Gogtay N, Nugent TF, Herman DH, Ordóñez A, Greenstein D, Hayashi KM, Clasen L, Toga AW, Giedd JN, Rapoport JL, Thompson PM. Dynamic Mapping of Normal Human Hippocampal Development. *Hippocampus*. 2006; 16:664–672. [PubMed: 16826559]
- Haier R, Benbow CP. Sex differences and lateralization in temporal lobe glucose metabolism during mathematical reasoning. *Dev. Neuropsychol*. 1995; 11:405–414.
- Haier RJ, Colom R, Schroeder D, Condon C, Tang C, Eaves E, Head K. Gray matter and intelligence factors: Is there a neuro-g? *Intelligence*. 2009; 37:136–144.
- Haier RJ, Jung RE, Yeo RA, et al. The neuroanatomy of general intelligence: sex matters. *NeuroImage*. 2005; 25:320–327. [PubMed: 15734366]
- Haier RJ, Siegel BV Jr, Nuechterlein KH, Hazlett E, Wu JC, Paek J, Browning HL, Buchsbaum MS. Cortical glucose metabolic rate correlates of abstract reasoning and attention studied with positron emission tomography. *Intelligence*. 1988; 12:199–121.
- Hanson JL, Chandra A, Wolfe BL, Pollak SD. Association between Income and the Hippocampus. *PLoS ONE*. 2011; 6(5):e18712. [PubMed: 21573231]
- Hines M. Gender Development and the Human Brain. *Annual Review of Neuroscience*. 2011; 34:69–88.

- Karama S, Colom R, Johnson W, Deary IJ, Haier RJ, Waber DP, Lepage C, Ganjavi H, Jung R, Evans AC. The Brain Development Cooperative Group. Cortical thickness correlates of cognitive performance accounted for by the general factor of intelligence in health children aged 6 to 18. *NeuroImage*. 2011; 55:1443–1453. [PubMed: 21241809]
- Lohman, DF. Complex information processing and intelligence. In: Sternberg, RJ., editor. *Handbook of intelligence*. Cambridge: Cambridge University Press; 2000. p. 285-340.
- Maguire EA, Gadian DG, Johnsrude IS, Good CD, Ashburner J, Frackowiak RSJ, Frith CD. Navigation-related structural change in the hippocampi of taxi drivers. *PNAS*. 2000; 97(8):4398–4403. [PubMed: 10716738]
- Martin-Loeches M, Bruner E, Cuétara JM, Colom R. Correlation between corpus callosum shape and cognitive performance in a sample of healthy young adults. *Brain Structure and Function*. (In Press).
- Morra JH, Tu Z, Apostolova LG, Green AE, Avedissian C, Madsen SK, Parikshak N, Hua X, Toga AW, Jack CR, Weiner MW, Thompson PM. The Alzheimer's Disease Neuroimaging Initiative. Validation of a fully automated 3D hippocampal segmentation method using subjects with Alzheimer's disease mild cognitive impairment, and elderly controls. *NeuroImage*. 2008; 43:59–68. [PubMed: 18675918]
- Nee DE, Jonides J. Dissociable contributions of prefrontal cortex and the hippocampus to short-term memory: Evidence for a 3-state model of memory. *NeuroImage*. 2011; 54:1540–1548. [PubMed: 20832478]
- Neubauer AC, Fink A. Intelligence and neural efficiency: Measures of brain activation versus measures of functional connectivity in the brain. *Intelligence*. 2009a; 37:223–229.
- Neubauer AC, Fink A. Intelligence and neural efficiency. *Neuroscience and biobehavioral Reviews*. 2009b
- Neubauer AC, Fink A, Schrausser DG. Intelligence and neural efficiency: the influence of task content and sex on the brain–IQ relationship. *Intelligence*. 2002; 30:515–536.
- Pantel J, O'Leary DS, Cretsingher K, Bockholt HJ, Keefe H, Magnotta VA, Andreasen NC. A New Method for the In Vivo Volumetric Measurement of the Human Hippocampus With High Neuroanatomical Accuracy. *Hippocampus*. 2000; 10:752–758. [PubMed: 11153720]
- Pessoa L. On the relationship between emotion and cognition. *Nature Reviews, Neuroscience*. 2008; 9:148–158.
- Piekema C, Kessels RPC, Rijpkema M, Fernandez G. The hippocampus supports encoding of between-domain associations within working memory. 2009 Article is online at <http://www.learnmem.org/cgi/doi/10.1101/lm.1283109>.
- Raven, J.; Raven, JC.; Court, JH. *Manual for Raven's Progressive Matrices and Vocabulary Scales*. San Antonio, TX: Pearson Assessment; 2004.
- Schmidt H, Jogia J, Fast K, Christodoulou T, Haldane M, Kumari V, Frangou S. No Gender Differences in Brain Activation During the N-Back Task: An fMRI Study in Healthy Individuals. *Human Brain Mapping*. 2009
- Shattuck DW, Leahy RM. BrainSuite: An Automated Cortical Surface Identification Tool. *Medical Image Analysis*. 2002; 8(2):129–142. [PubMed: 12045000]
- Sheppard LD, Vernon PA. Intelligence and speed of information-processing: A review of 50 years of research. *Personality and Individual Differences*. 2008; 44:535–551.
- Squire LR, van der Horst AS, McDuff SGR, Frascino JC, Hopkins RO, Mauldin KN. Role of the hippocampus in remembering the past and imagining the future. *PNAS*. 2010; 107(44):19044–19048. [PubMed: 20956323]
- Tang CY, Eaves EL, Ng JC, Carpenter DM, Kanellopoulou I, Mai X, Schroeder DH, Condon CA, Colom R, Haier RJ. Brain networks for working memory and factors of intelligence assessed in males and females with fMRI and DTI. *Intelligence*. 2010; 38:293–303.
- Thurstone L. *Primary mental abilities*. Psychometric Monographs. 1938; 1
- Yela, M. *Rotación de Figuras Macizas [Rotation of Solid Figures]*. Madrid, TEA: 1969.

### Highlights

- Automated segmentation of the hippocampus was applied
- 21 cognitive measures were administered to 104 young adults
- Hippocampal structural differences were related to spatial intelligence, working memory, and executive control
- Sex interacted with general intelligence (*g*), along with abstract and spatial intelligence
- Verbal intelligence, attention, and processing speed were not related to the hippocampus

A

**Abstract-Fluid Intelligence**

**APM**

**DAT-AR**

**DAT-AR**

**PMA-R**

abcabcabc...      cabcx

**Verbal-Crystallized Intelligence**

**DAT-VR**

... is to water like eating is to ...

- (a) Travelling-driving
- (b) Foot-enemy
- (c) Drinking-bread
- (d) Girl-industry
- (e) Drinking-enemy

**PMA-V**

Stout >>> Sick – Fat – Short – Rude

**DAT-VR**

Which number must be substituted by P if the expression  $5P + 2 = 58$  is correct?

- (a) 3
- (b) 4
- (c) 7
- (d) 9
- (e) None of them

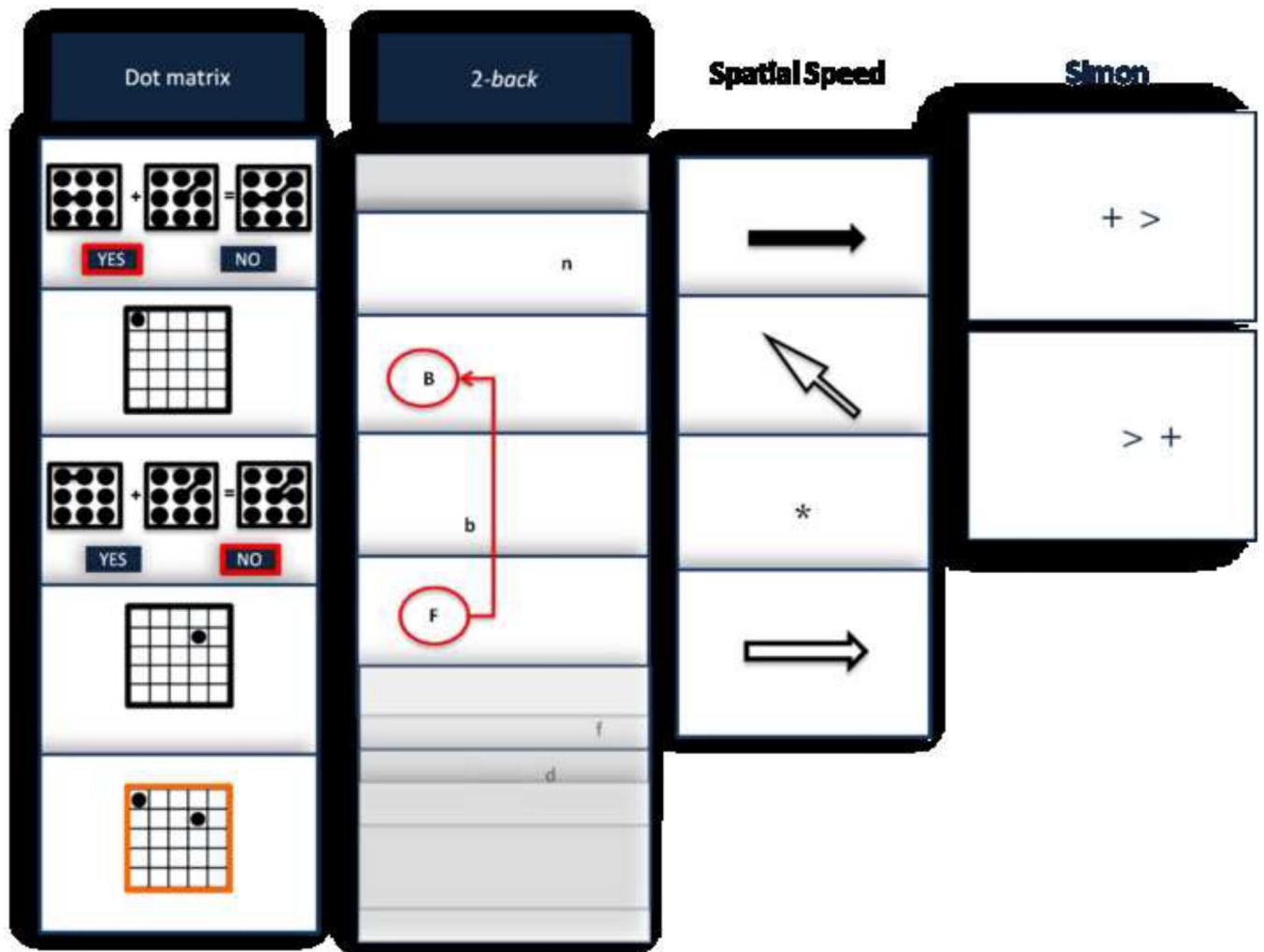
**Spatial Intelligence**

**DAT-SR**

**PMA-S**

**ROTATION OF SOLID FIGURES**

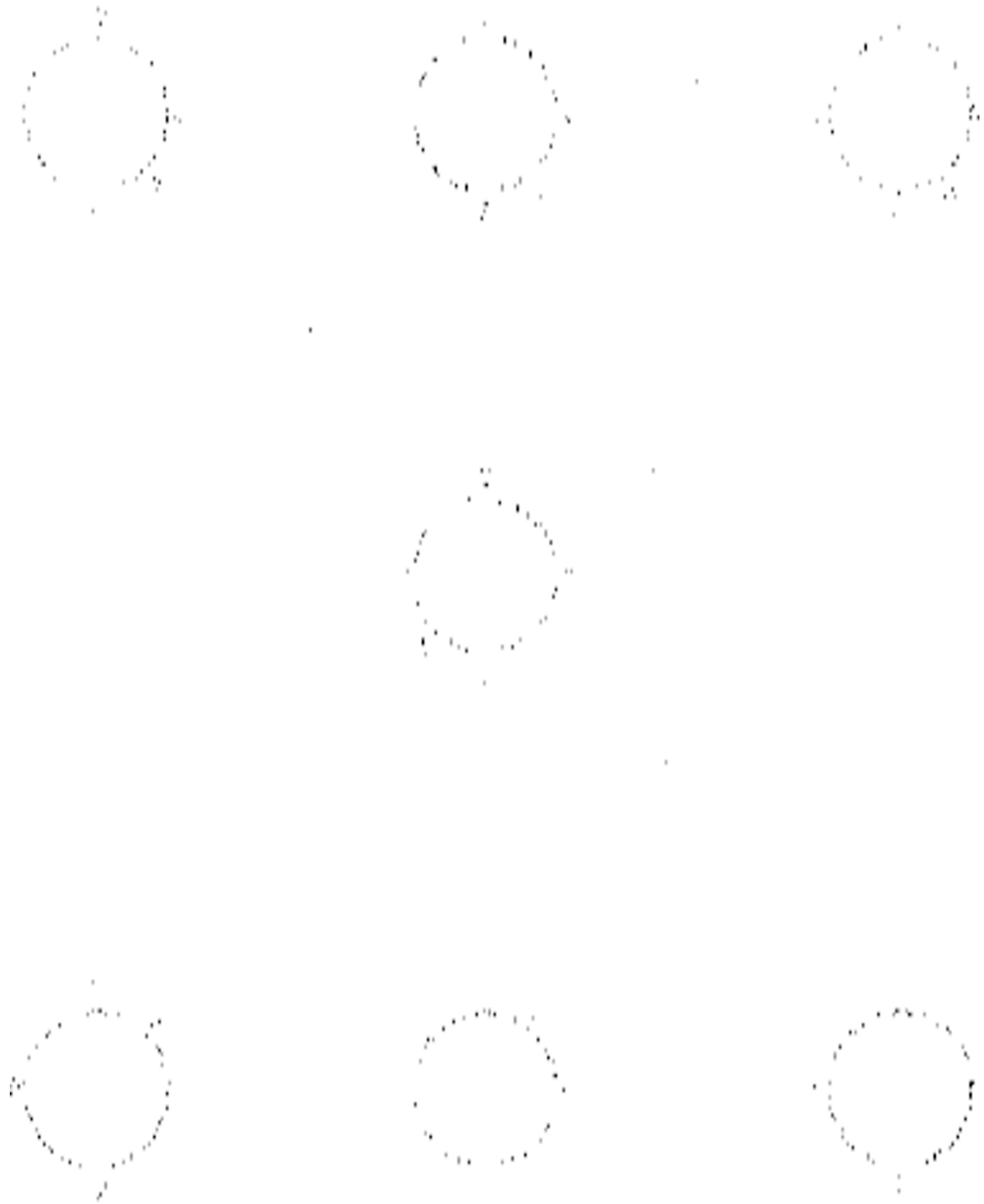
B



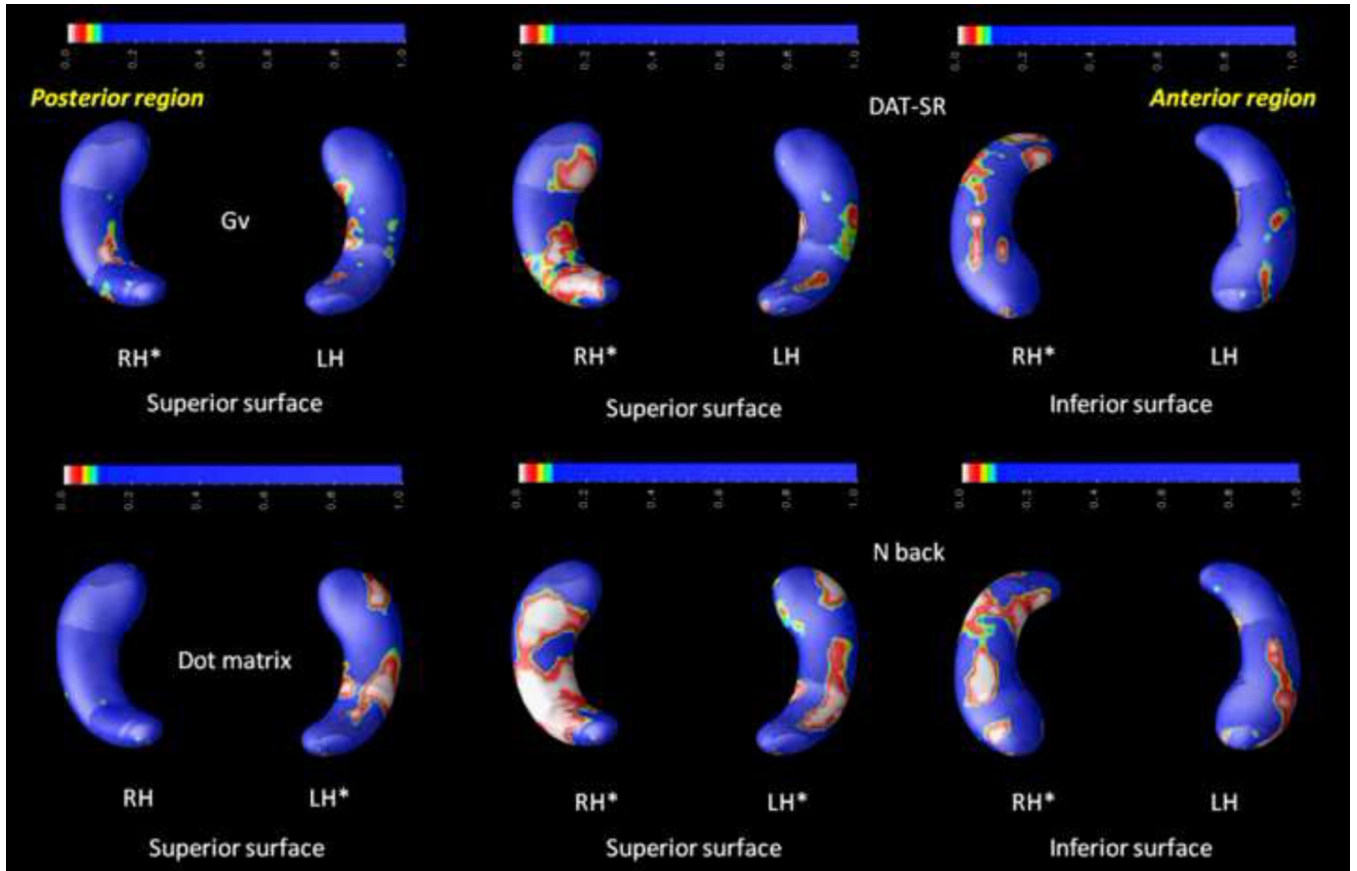
**FIGURE 1.**

Examples of tests and tasks. Top panel depicts items from the intelligence tests, whereas bottom panel shows some captions for working memory (dot matrix), executive control (n back), processing speed (spatial recognition speed), and attention (Simon).



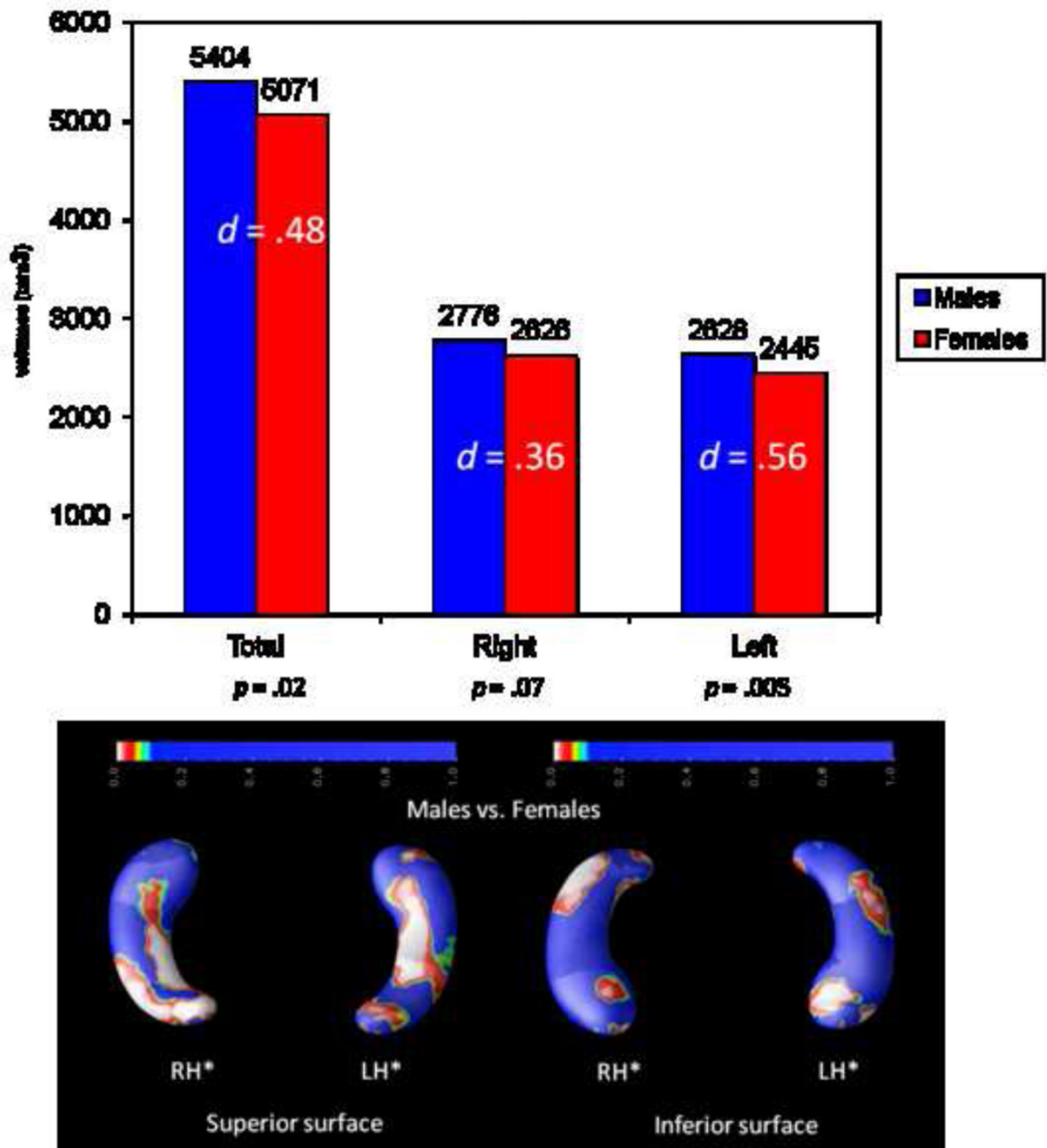


**FIGURE 2.** SEM model for the cognitive factors. Gf = abstract-fluid intelligence, Gc = verbal-crystallized intelligence, Gv = spatial intelligence, WM = working memory, UPD = executive updating, Speed = processing speed, ATT = attention



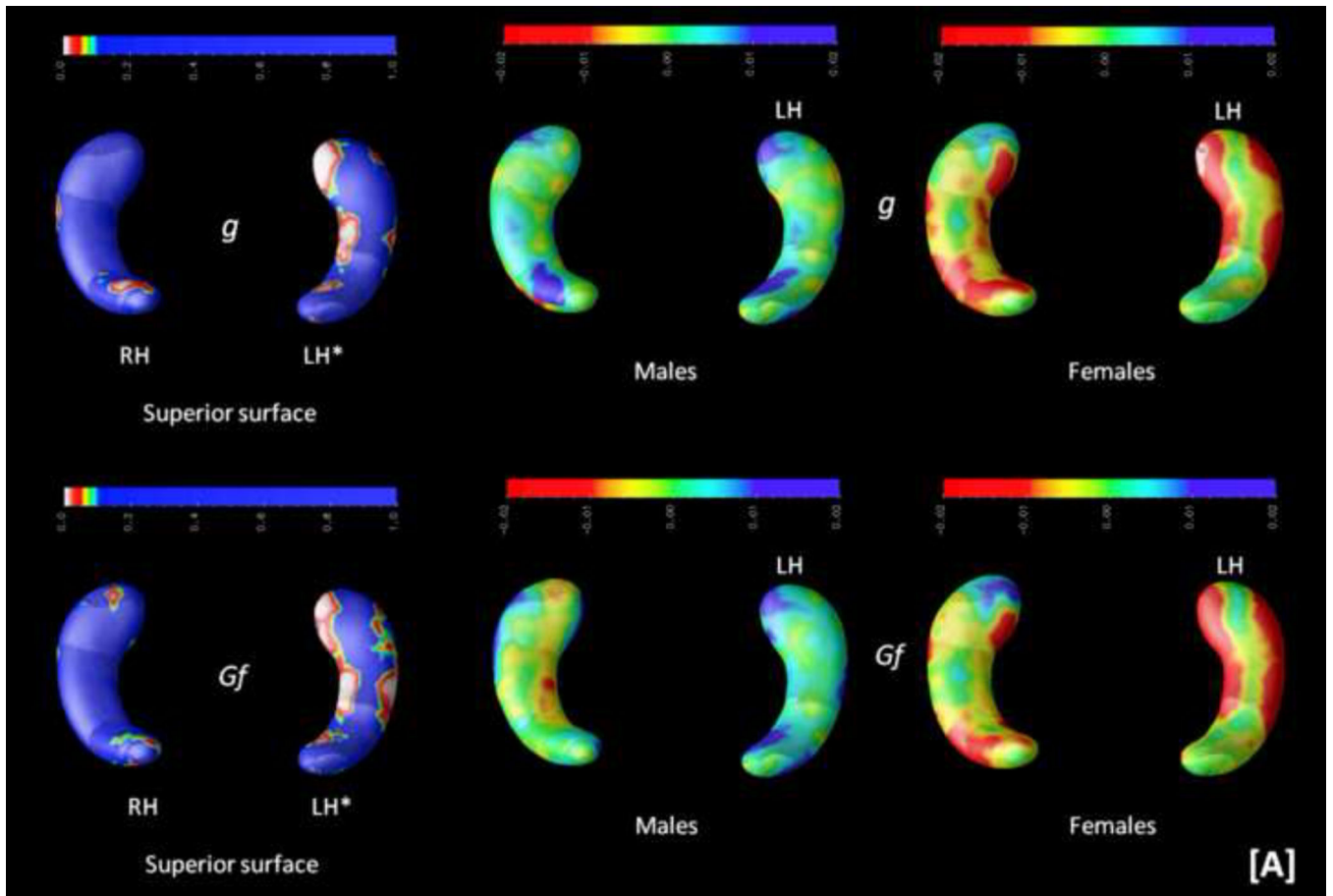
**FIGURE 3.**

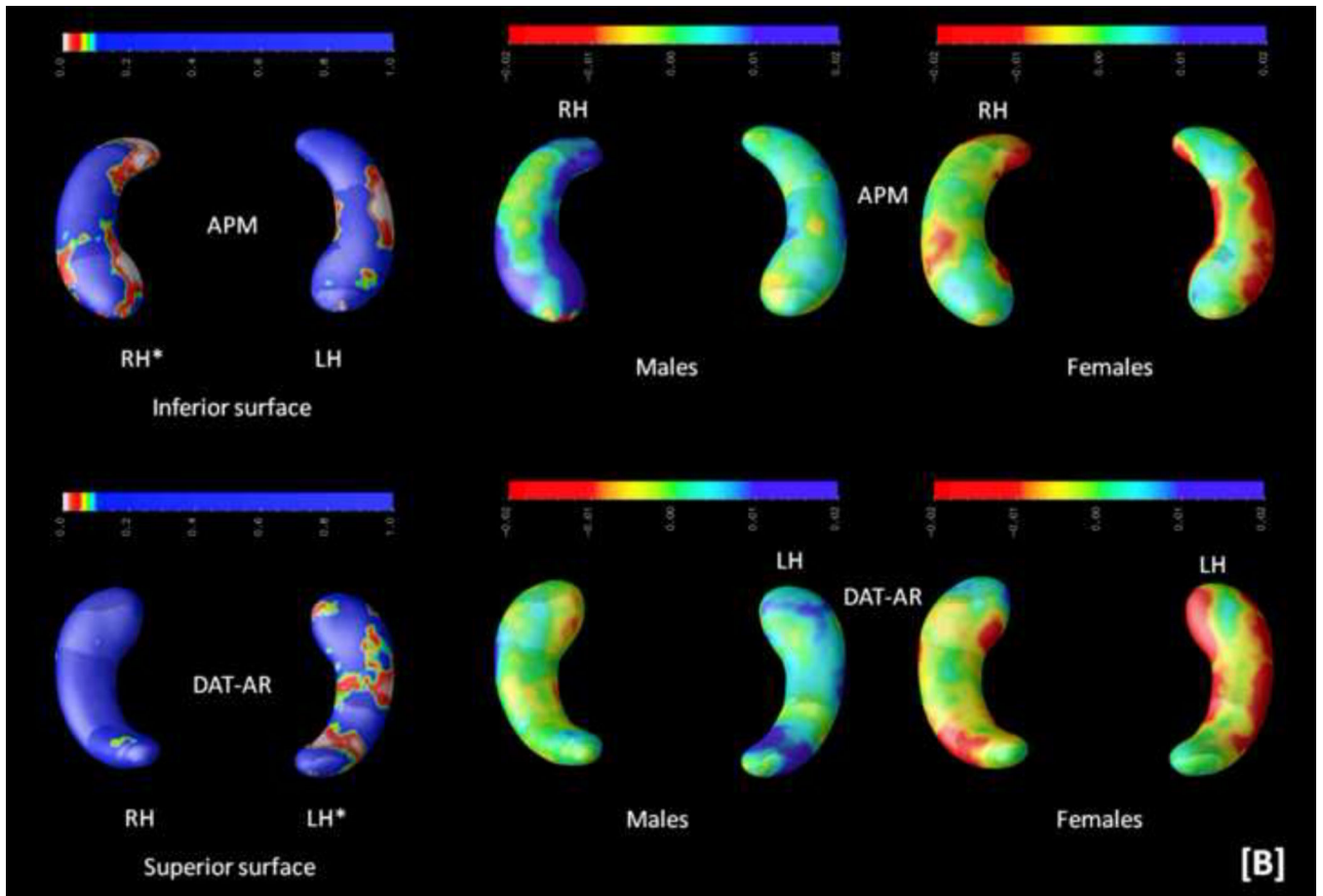
Probability maps (*p*-maps) for the complete group ( $N = 104$ ) controlling for sex, age, and handedness. Significant results were found for general spatial intelligence (Gv), the spatial relations (SR) subtest from the DAT battery (DAT-SR), the spatial working memory task (dot matrix), and the spatial executive control task (n back). Asterisks (\*) highlight significant effects ( $p < .05$ , corrected for multiple comparisons). Radiological convention is the framework for depicting RH (right hippocampus, on the readers' left) and LH (left hippocampus, on the reader's right) along with views for superior and inferior surfaces. The two latter views are represented only when significant findings are found for both.

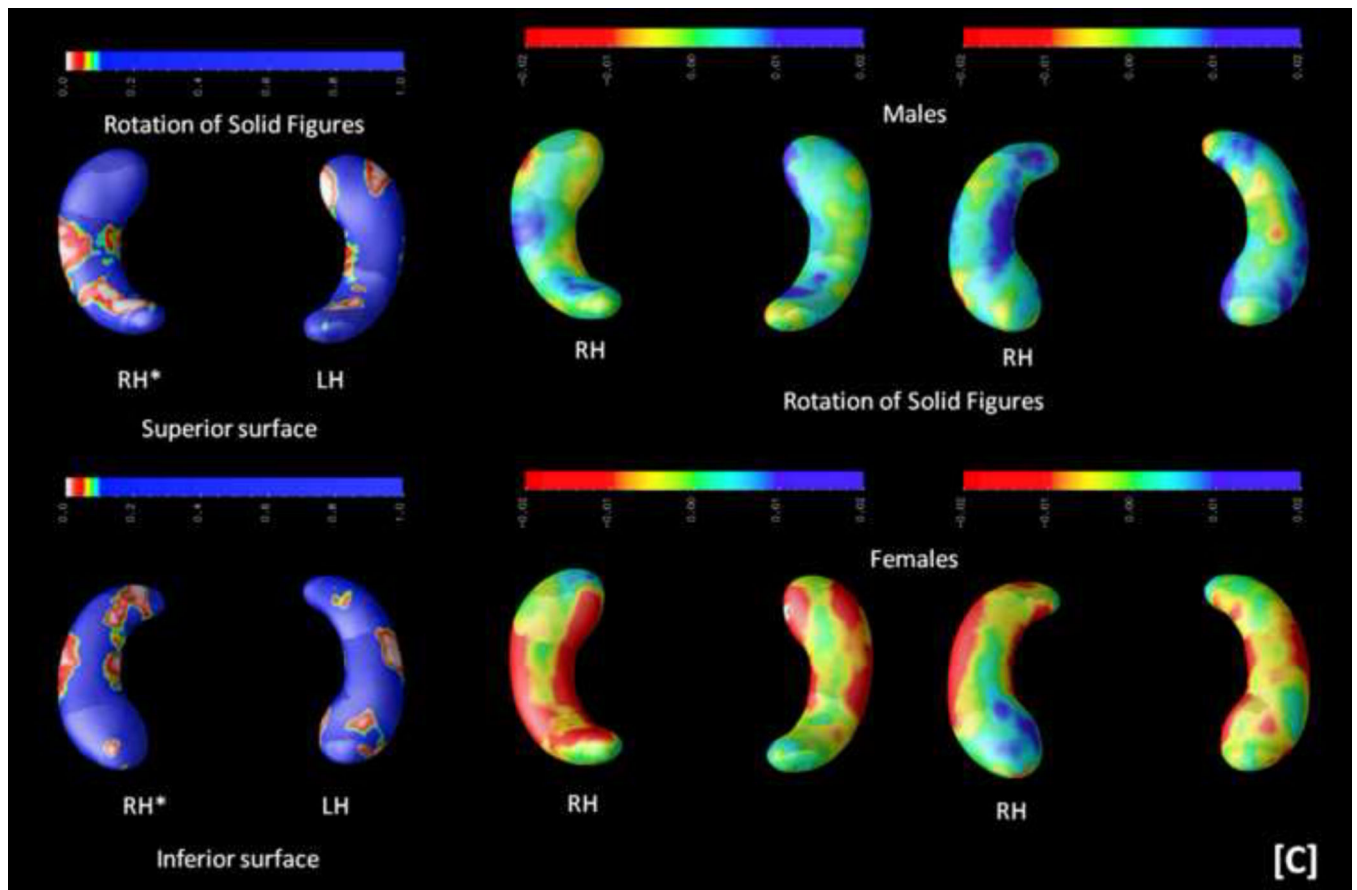


**FIGURE 4.**

Sex differences in total hippocampal volumes (top panel, effect sizes --*d*-- are also shown) and regional structural differences between males and females (bottom panel). Asterisks (\*) highlight significant effects ( $p < .05$ , corrected). Radiological convention is the framework for depicting RH (right hippocampus, on the readers' left) and LH (left hippocampus, on the reader's right) along with views for superior and inferior surfaces.







**FIGURE 5.**

Probability maps ( $p$ -maps) showing the interaction among sex, cognition, and hippocampal structure are depicted on the left. Beta maps indicating the direction of these interactions are shown on the right (yellow and red depict negative values, whereas blue depicts positive values). Significant findings were observed for [A] the general factor of intelligence ( $g$ ) and abstract-fluid intelligence ( $Gf$ ), [B] the Advanced Progressive Matrices (APM) test and the abstract reasoning (AR) subtests from the DAT battery (DAT-AR), and [C] the rotation of solid figures test. Asterisks (\*) highlight significant effects ( $p < .05$ , corrected). Radiological convention is the framework for depicting RH (right hippocampus, on the readers' left) and LH (left hippocampus, on the reader's right) along with views for superior and inferior surfaces. These two latter views are represented only when significant findings are found for both.

TABLE 1

Means and SDs (standard deviation) for males and females on the cognitive factors and measures; the effect size (*d*) is also shown

Variable	Males (N = 45)		Females (N = 59)		<i>d</i>
	Mean	SD	Mean	SD	
<i>g</i> (General intelligence)	159	25	155	23	0.16
Gf (Fluid intelligence)	45	8	46	8	-0.21
Gc (Crystallized intelligence)	60	10	57	10	0.38
Gv (Spatial intelligence)	54	15	52	14	0.13
WMC (Working memory)	215	17	211	19	0.22
Executive control	50	9	51	9	-0.17
Attention	551	65	550	55	-0.02
Processing Speed	717	176	729	156	-0.07
RAPM	12	3	12	2	-0.06
PMA-R*	18	4	20	4	-0.49
DAT-AR	15	3	14	4	0.17
PMA-V	34	6	32	7	0.34
DAT-VR	14	4	13	3	0.17
DAT-NR	12	3	12	3	0
Rotation of Solid Figures**	10	4	8	3	0.64
PMA-S	28	10	27	10	0.10
DAT-SR	15	5	17	4	-0.31
Reading Span	98	5	98	8	0.10
Computation Span	61	14	58	12	0.24
Dot Matrix	55	5	55	5	0
2 Back	17	5	16	6	0.17
Letter memory	17	4	18	3	-0.04
Keep track*	15	5	18	4	-0.52
Verbal attention	570	76	549	68	-0.29
Numerical attention*	613	105	605	69	-0.08
Spatial attention	469	65	494	54	0.43

Variable	Males (N = 45)		Females (N = 59)		<i>d</i>
	Mean	SD	Mean	SD	
Verbal speed	668	212	613	132	-0.32
Numerical speed	860	305	936	272	0.26
Spatial speed	623	175	639	143	0.10

\*  $p < .05$ ,

\*\*  $p < .01$