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

Roberto Colom a,b,⁎, Jason L. Stein c, Priya Rajagopalan c, Kenia Martínez a,b, David Hermel c, Yalin Wang c, Juan Álvarez-Linera d, Miguel Burgaleta a,b, Mª. Ángeles Quiroga e, Pei Chun Shih a,b, Paul M. Thompson c

a Facultad de Psicología, Universidad Autónoma de Madrid, 28049 Madrid, Spain
b Fundación CIEN, Fundación Reina Sofia, C/Valderrebollo, 5, 28031 Madrid, Spain
c Laboratory of Neuroimaging (LONI), UCLA School of Medicine, 635 Charles E. Young Drive South, Suite C255, Los Angeles, CA 90095-7334, USA
d Hospital Ruber Internacional, Fundación CIEN, Fundación Reina Sofia, C/Valderrebollo, 5, 28031 Madrid, Spain
e Facultad de Psicología, Universidad Complutense de Madrid, Campus de Somosaguas, 28223 Pozuelo de Alarcón, Madrid, Spain

1. 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. In addition, 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, Kessels, Rijpkema, & Fernandez, 2009). The hippocampus contributes to a tempo-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). In addition, 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 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, Karama, Jung, & Haier, 2010; Colom & Thompson, 2011; Colom et al., 2009; Colom et al., 2010; 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, Rebollo, Abad, & Shih, 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. Fig. 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.

2. Method

2.1. 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).\(^1\) 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 3 T scanner.\(^2\) The scans were lineely registered to the ICBM template with a 9 parameter transformation (3 translations, 3 rotations, and 3 scales) using the Mitracc algorithm. Aligned images were resampled in an isotropic space with a voxel size of 1 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, overlap = .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,\(^1\) 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 subsets. Bruner, Cuétara, Colom, and Martin-Loeffes (2012, 2010); Bruner, Martin-Loeches, and Colom (2010); Bruner, Martin-Loeches, Burgaleta, and Colom (2011), and Martin-Loeches, Bruner, Cuétara, and Colom (in press) based their research in 2D midsagittal brain shape variations.

\(^2\) GEHC Waukesha, WI, 3 T 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 1.5 mm thickness, full brain coverage (220 slices), matrix 266 × 266 FDOV 24 (isotropic voxels 7.3 cm\(^3\)).
Fig. 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).
as detailed below. A second Pipeline was then applied for
to obtain 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 delini-
eation of the hippocampus, the automatic recognition algo-

2.2. 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 represen-
tative scores.

Abstract-fluid intelligence (Gf) was measured by the Ad-
anced 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 3×3
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 com-
promises 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 adminis-
tered (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 quantita-
tive 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, Karama
et al., 2010; Colom, Quiroga, 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 × 3 trials = 12 trials total). The computation span
task includes a verification task and a recall task. Six s is 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 × 3 trials
each = 15 trials total). In the dot matrix task, a matrix equation
must be verified and a dot location displayed in a 5×5 grid
must be retained. The matrix equation is presented during a
maximum of 4.5 s 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 × 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, Abad
Quiroga, Shih, & Flores-Mendoza, 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
core 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, low-
erase 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, and 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, Karama et al., 2010; Colom, Quiroga, 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 × 30 trials each = 60 trials total). For example [Verbal speed]: [Memory Set] [First screen] A[b] [Second screen] A[c] [Answer screen, Same meaning?] ‘Yes’. 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.

2.3. Manual delineation of the hippocampus

The training brains set comprised twenty subjects (half were females) 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).

Table 1

<table>
<thead>
<tr>
<th>Variable</th>
<th>Males (N=45)</th>
<th>Females (N=59)</th>
<th>t</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>g (General intelligence)</td>
<td>159</td>
<td>155</td>
<td>0.16</td>
<td>0.6</td>
</tr>
<tr>
<td>GF (Fluid intelligence)</td>
<td>45</td>
<td>46</td>
<td>0.21</td>
<td>0.17</td>
</tr>
<tr>
<td>GC (Crystallized intelligence)</td>
<td>60</td>
<td>57</td>
<td>0.38</td>
<td>0.18</td>
</tr>
<tr>
<td>GV (Spatial intelligence)</td>
<td>54</td>
<td>52</td>
<td>0.13</td>
<td>0.19</td>
</tr>
<tr>
<td>WMC (Working memory)</td>
<td>215</td>
<td>211</td>
<td>0.22</td>
<td>0.10</td>
</tr>
<tr>
<td>Executive control</td>
<td>50</td>
<td>51</td>
<td>0.17</td>
<td>0.11</td>
</tr>
<tr>
<td>Attention</td>
<td>551</td>
<td>550</td>
<td>0.02</td>
<td>0.12</td>
</tr>
</tbody>
</table>

...
Fig. 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, and ATT = attention.

Fig. 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.
2.3.1. 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.

2.3.2. 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 color 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

![Fig. 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.](http://dx.doi.org/10.1016/j.intell.2013.01.002)
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.

3. 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 modeling (SEM) — Fig. 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 Fig. 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 Fig. 3.
significant relationships were mainly found in posterior areas for
general spatial intelligence (Gv), 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 Fig. 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 concentrat-
ed on the posterior regions of the right hippocampus, but they
are more widespread on the left side.

Third, because of the observed sex differences, interac-
tions among hippocampal structure, cognitive measures, and
sex were computed (controlling for age and handedness).

Results can be inspected in Fig. 5.

Significant findings were observed for the general factor of
intelligence (g), abstract-fluid intelligence (Gf), 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 Fig. 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.

4. Discussion

Here we report findings providing what might be consid-
ered an interesting general picture. Firstly, varied — but not
general — measures comprising spatial processing require-
ments are significantly related to hippocampal structural
differences. Specifically, spatial intelligence, spatial executive
control processes, and spatial working memory. It is notewor-
thy that abstract processing, verbal processing, 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 intelligences — as well as abstract-fluid intelligence
(Gf) 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 sig-
nificant associations were mainly positive for males but negative
for females, suggesting higher efficiency for women (Haier et al.,
1988; Neubauer & Fink, 2009a). 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 (2009b), Haier and Benbow
(1995) failed to find positive evidence for the neural efficiency
hypothesis, but they reported substantial sex differences.

Neubauer, Fink, and Schrausser (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 (2009b), functional results
may or may not fit structural ones. At this structural level,
females might show greater efficiency requiring less neural
material for achieving behavioral 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., 2008). Hippocam-
pus 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, Chandra,
Wolfe, & Pollak, 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, Jung, Yeo, 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
especially 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 perfor-
ance 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 recognized – but frequently neglected – APM spatial requirements. (Ackerman, Beier, & Boyle, 2002).

Acknowledgments

This work was supported by grants SEJ-2006-07890 (Ministry of Education and Culture, Spain), PR2008-0038, and PSI2010-20364 (Ministry of Science and Innovation, Spain).

Software for automated segmentation was developed at LONI (Laboratory of Neuroimaging), School of Medicine, UCLA.

References


