Human hippocampus associates information in memory

KATHARINA HENKE*†‡, BRUNO WEBER*, STEFAN KNEIFEL†, HEINZ GREGOR WIESER*, AND ALFRED BUCK†

*Department of Neurology and †PET Center, University Hospital Zürich, 8091 Zürich, Switzerland

Edited by Larry R. Squire, University of California, San Diego, CA, and approved February 19, 1999 (received for review November 2, 1998)

ABSTRACT The hippocampal formation, one of the most complex and vulnerable brain structures, is recognized as a crucial brain area subserving human long-term memory. Yet, its specific functions in memory are controversial. Recent experimental results suggest that the hippocampal contribution to human memory is limited to episodic memory, novelty detection, semantic (deep) processing of information, and spatial memory. We measured the regional cerebral blood flow by positron-emission tomography while healthy volunteers learned pairs of words with different learning strategies. These led to different forms of learning, allowing us to test the degree to which they challenge hippocampal function. Neither novelty detection nor depth of processing activated the hippocampal formation as much as semantically associating the primarily unrelated words in memory. This is compelling evidence for another function of the human hippocampal formation in memory: establishing semantic associations.

The discovery that the mediotemporal brain regions, particularly the hippocampal formations, are essential for human memory (1, 2) set the ground for neuroscientific theories and experimental practice during the past 40 years. In the years following this discovery, research with amnesic patients led to the finding that memory is not a unitary system but is divided into subsystems, each supported by a different but partially overlapping neuronal network. The function of the human hippocampal formation was pinned down to declarative memory alone and became specified even further to declarative learning/consolidation (3, 4), episodic memory (5), novelty detection (6–9), the retrieval of deeply encoded items (10), and spatial learning (11–13). At the same time, learning experiments with rats (14–17) indicated that the hippocampal formation is important for the establishment of associations between components of episodes in memory. The experience of an episode typically involves the simultaneous processing of diverse sensory inputs, bodily sensations, thoughts, and emotions in distributed cortical regions, creating patterns of co-activations in the cortex. The composition of these coactivations needs to be stored in memory for the later recovery of some or all aspects of that episode. The anatomy and physiology of the hippocampal formation (dentate gyrus, Ammon’s horn/hippocampus proper, presubiculum) lends itself to store such patterns of neuronal coactivations temporarily (3, 4, 14–24). The less complex a scene is, the fewer associations are required to memorize it, and, thus, the less hippocampal activation can be expected. We tested this hypothesis in the human, reducing the complexity of a natural scene to the sensorily restricted environment of a word-learning experiment in the positron-emission tomography (PET) scanner. Visually presented pairs of semantically unrelated nouns had to be learned by one of three strategies inducing a (i) deep or (ii) shallow encoding of the words in isolation of each other or inducing (iii) both the deep encoding of the words as separate entities and the creation/storage of semantic associations between the two words. It was our hypothesis that the latter task would activate the hippocampal formation more than either of the single-item learning tasks. We also expected that the magnitude of hippocampal activation underlying the associative process would exceed that underlying novelty detection, which was induced by a further manipulation.

METHODS

Task Design. The experiment consisted of the following tasks. Associative Word Learning (AWL) required subjects to decide whether or not the two presented words fit together in meaning. Subjects answered with their right hand, pressing the left key of a computer mouse for “fit” and the right key for “no fit” (Fig. 1). This task induces a comparison of the meanings of the two abstract nouns on grounds of subjective criteria and thereby leads to the storage of the two words in relation to each other. Associative Word Learning—Old (AWLO) was identical except that the presented words had been studied immediately preceding scanning. For this preceding run, the 40-word stimuli were presented as a list of eight lines, each line consisting of five words. The two words that were going to be presented as a pair during the subsequent associative learning task were not presented side by side in this list. The subject’s instruction was to read through the list of words and to indicate whether there are words that he hardly ever thinks/speaks/reads/hears and whose meanings, therefore, are less clear. The list was presented for as long as the subject needed it. This instruction induces a deep processing of the individual words, whereby they become contextually old for the following associative learning task. Deep Single Word Encoding (DSWE) required subjects to decide for each of the two presented words whether it was pleasant or unpleasant. Subjects pressed the left key if one of the two words was pleasant and the right if both or none was pleasant. This task leads to a semantic evaluation (deep encoding) (25) of each word individually. For Shallow Single Word Encoding (SSWE), subjects counted the vowels and added them over both words. The left key was used if the sum of the vowels across both words was less than 6, and the right key was used if the sum was equal or greater than 6. This task induces a shallow encoding of words, because words are processed with respect to their appearance rather than their meaning (25). Three minutes after the completion of each learning task, a retrieval of the learned word pairs was carried out without PET scan. The retrieval task was always the same: the words presented during the previous scan were shown again, either in the same combination as before or recombined to form new pairs. The task was to decide which combinations were old/new. Subjects used the left key for “old combination” and the right for “new combination” (Fig. 1). Subjects prac-

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. §1734 solely to indicate this fact.

PNAS is available online at www.pnas.org.

5884
PET Scanning. PET scans were acquired on a whole-body scanner (Advance; GE Medical Systems, Waukesha, WI) in three-dimensional mode with a 15-cm axial field of view. For each scan, 400–450 MBq H$_2^{15}$O was administered as a slow bolus with a remotely controlled injection device. PET counts were recorded over 60 s after the arrival of the bolus in the brain. Attenuation-corrected data were reconstructed into 35 image planes. The accumulated radioactivity counts over 60 s were taken as a measure for cerebral blood flow. Statistical parametric mapping was performed in the following way. First, head movements between scans were corrected by using the least-squares method implemented in SPM96 (27, 28). For the transformation into stereotaxic space (29) the images of each subject were summed and subsequently transformed by using an affine spatial normalization (SPM96). The normalization included linear transformations and nonlinear basis function-based deformations. The resulting transformation matrix subsequently was used to transform each individual scan. To ameliorate residual anatomical differences after spatial normalization the scans were smoothed with a Gaussian filter (15 mm full width at half maximum). Global effects such as varying injected activities were removed by dividing each voxel value by the global mean of gray matter voxels. The difference between conditions then was evaluated voxel by voxel, using t statistics, which then were transformed into normally distributed z statistics.

For the region-of-interest (ROI) analysis, we used data that had been realigned, corrected for global effects, spatially normalized, and smoothed. Circular regions of interest were placed on the slices of an MRI, which also was normalized into Talairach space. These ROIs then were applied to the PET data by software developed in our laboratory (30). For the first ROI analysis, the mean activity associated with a task was determined per subject and ROI and then was used for the statistical analysis. For the second ROI analysis, differences between mean activations associated with two tasks were computed for each subject and ROI. These difference activations were tested for significance with t tests.

RESULTS

Cognitive Performance. The direction of reaction time (RT) differences between tasks indicates that task difficulty (expressed in RT) was not positively related to the degree of hippocampal blood flow during tasks (learning tasks: $t_{24} = 1.45$, $P = 0.06$; $t_{24} = 1.67$, $P = 0.00$ between AWL and AWLO; $t_{24} = 4.24$, $P = 0.00$). The retrieval performance differed significantly between AWL and SSWE ($t_{24} = 2.49$, $P = 0.01$, one-tailed), between AWL and SSWE ($t_{24} = 2.86$, $P = 0.00$), as well as between AWLO and DSWE ($t_{24} = 4.74$, $P = 0.00$), between AWLO and SSWE ($t_{24} = 4.58$, $P = 0.00$), and also between AWL and AWLO ($t_{24} = 1.9$, $P = 0.04$).

Two pilot studies with 10 subjects in each had shown that single-word recognition (list of 40 words: 20 targets, 20 distractors) was as good after DSWE ($M = 64\%$, $SD = 17\%$) as after AWL ($M = 67\%$, $SD = 15\%$; $t_9 = 0.53$, $P = 0.61$, two-tailed). This indicates that depth of single-word processing is equal in these two tasks. Single-word recognition after SSWE ($M = 19\%$, $SD = 29\%$) was, however, significantly

---

\*Percentage correct was computed by subtracting the percentage of false alarms from the percentage of hits.
Cerebral Blood Flow. Pairwise subtraction analysis of the intersubject-averaged PET images was performed (27, 28) (Table 1). Because we were particularly interested in medio-temporal blood flow changes, the discussion will be limited to this area. The comparison of AWL with SSWE indicates blood flow differences in areas subserving depth of processing and associative processing. This comparison revealed significant blood flow increases in the right hippocampal formation and amygdala in addition to other locations (Table 1 and Fig. 2). The comparison of AWL with DSWE is the crucial comparison to test the “association” hypothesis, because it reveals blood flow changes underlying associative processing alone. There was significantly elevated blood flow in the same anterior hippocampal area as in the first contrast, besides other locations (Table 1 and Fig. 2). Some of these other locations are not subsumed in the list of peak activation differences resulting from the comparison between AWL and SSWE (Table 1). There may be several reasons for this. Activation differences between AWL and DSWE may have been due to blood flow decreases during DSWE (but not SSWE), or they may not have been large enough to survive in the comparison of AWL with SSWE, which yields more areas with activation differences and greater activation differences than AWL/DSWE. Furthermore, SSWE might have increased the blood flow in regions that AWL does (but not DSWE), thereby subtracting activation out in AWL/SSWE. Finally, there may be regions with significant blood flow differences that are part of a larger area with significant blood flow differences. Only the voxel with the peak difference within this larger area gets indexed in the SPM list and in Table 1. An example for this is the significant activation difference in the left fusiform gyrus resulting from AWL/DSWE. This activation difference is also present in the comparison of AWL with SSWE, but it does not get indexed because it is part of a wider-spread activation in AWL/SSWE.

We also computed the contrast between AWLO and DSWE, which is the most rigorous test for the “association” hypothesis, because the words presented during associative learning were contextually old (no novelty effect) and had been deeply encoded already, whereas those in DSWE were novel. Therefore, hippocampal activations may get subtracted out by activations in the same region underlying novelty detection and/or deep encoding, which are part of the reference task. This was, however, not the case: significant blood flow increases were localized in the left parahippocampal gyrus and hippocampus as well as other locations (Table 1).

The comparison between AWL and AWLO reveals blood flow underlying the processing of contextually novel versus old words. There was no significant hippocampal activation resulting from this contrast. DSWE compared with SSWE yields blood flow changes associated with depth of processing. This contrast also indicated no significant hippocampal activation (Table 1). Thus, neither novelty detection nor depth of processing, but the associative process alone, significantly increased the hippocampal blood flow in this whole-brain analysis.

To further explore task-induced differences in hippocampal blood flow, two ROI analyses were conducted. The first was guided by the result of the subtraction analyses, indicating that the voxel with the highest blood flow difference between AWL and DSWE and AWL and SSWE is located in the pes of the right hippocampal formation. The aim of the first ROI analysis was to analyze task-induced activation differences with the mean activity per subject in the entire pes of hippocampus (not worse than single-word recognition after DSWE [(15) = 3.95, P = 0.00] and single-word recognition after AWL [(14) = 4.39, P = 0.00].

### Table 1. Peaks of blood flow difference

<table>
<thead>
<tr>
<th>Contrast</th>
<th>Isolated cognitive processes</th>
<th>Brain area</th>
<th>Coordinates</th>
<th>z value</th>
</tr>
</thead>
<tbody>
<tr>
<td>AWL/SSWE</td>
<td>Deep single-word encoding and associating words</td>
<td>L. inferior frontal gyrus (BA 45/47)</td>
<td>38, 26, -8</td>
<td>5.50</td>
</tr>
<tr>
<td></td>
<td></td>
<td>R. cerebellum</td>
<td>32, -74, -44</td>
<td>5.27</td>
</tr>
<tr>
<td></td>
<td></td>
<td>R. superior temporal gyrus (BA 38)</td>
<td>34, 6, -16</td>
<td>4.15</td>
</tr>
<tr>
<td></td>
<td></td>
<td>R. hippocampus</td>
<td>24, -12, -20</td>
<td>3.94</td>
</tr>
<tr>
<td></td>
<td></td>
<td>R. amygdala</td>
<td>26, -4, -16</td>
<td>3.63</td>
</tr>
<tr>
<td></td>
<td></td>
<td>R. cuneus/precuneus</td>
<td>4, -70, 20</td>
<td>3.55</td>
</tr>
<tr>
<td></td>
<td></td>
<td>R. anterior cingulate</td>
<td>6, 28, 8</td>
<td>3.55</td>
</tr>
<tr>
<td>AWL/DSWE</td>
<td>Associating words</td>
<td>R. inferior parietal gyrus (BA 40)</td>
<td>56, -32, 56</td>
<td>3.75</td>
</tr>
<tr>
<td></td>
<td></td>
<td>L. fusiform gyrus (BA 20)</td>
<td>-30, -34, -24</td>
<td>3.34</td>
</tr>
<tr>
<td></td>
<td></td>
<td>L. postcentral gyrus (BA 1/2)</td>
<td>-60, -14, 24</td>
<td>3.32</td>
</tr>
<tr>
<td></td>
<td></td>
<td>L. precentral gyrus (BA 4/6)</td>
<td>-16, -22, 52</td>
<td>3.23</td>
</tr>
<tr>
<td></td>
<td></td>
<td>L. anterior cingulate (BA 24)</td>
<td>-12, 22, -4</td>
<td>3.22</td>
</tr>
<tr>
<td></td>
<td></td>
<td>R. hippocampus</td>
<td>22, -10, -20</td>
<td>3.01</td>
</tr>
<tr>
<td>AWLO/DSWE</td>
<td>Associating words</td>
<td>L. parahippocampal gyrus (BA 36)</td>
<td>-30, -30, -24</td>
<td>3.49</td>
</tr>
<tr>
<td></td>
<td></td>
<td>L. uncus</td>
<td>-32, -22, -28</td>
<td>3.29</td>
</tr>
<tr>
<td></td>
<td></td>
<td>L. globus pallidus</td>
<td>-24, -6, -8</td>
<td>3.25</td>
</tr>
<tr>
<td></td>
<td></td>
<td>R. precentral gyrus (BA 6)</td>
<td>48, -8, 28</td>
<td>3.48</td>
</tr>
<tr>
<td></td>
<td></td>
<td>R. middle temporal gyrus (BA 21)</td>
<td>60, 8, -16</td>
<td>3.30</td>
</tr>
<tr>
<td></td>
<td></td>
<td>R. superior temporal gyrus (BA 38)</td>
<td>54, 12, -12</td>
<td>3.25</td>
</tr>
<tr>
<td></td>
<td></td>
<td>R. medial frontal gyrus</td>
<td>12, -4, 56</td>
<td>3.26</td>
</tr>
<tr>
<td></td>
<td></td>
<td>R. cingulum at the level of BA 24</td>
<td>22, 14, 28</td>
<td>3.51</td>
</tr>
<tr>
<td></td>
<td></td>
<td>L. inferior frontal gyrus (BA 44/45)</td>
<td>-52, 20, 8</td>
<td>5.78</td>
</tr>
<tr>
<td></td>
<td></td>
<td>R. cerebellum</td>
<td>38, -76, -48</td>
<td>4.82</td>
</tr>
<tr>
<td></td>
<td></td>
<td>R. middle temporal gyrus (BA 21)</td>
<td>66, -6, -24</td>
<td>4.19</td>
</tr>
<tr>
<td></td>
<td></td>
<td>R. superior temporal gyrus (BA 38)</td>
<td>42, 8, -24</td>
<td>4.07</td>
</tr>
<tr>
<td></td>
<td></td>
<td>R. middle frontal gyrus (BA 8/9)</td>
<td>52, 20, 48</td>
<td>3.33</td>
</tr>
<tr>
<td></td>
<td></td>
<td>L. caudate nucleus</td>
<td>-10, 16, 12</td>
<td>3.77</td>
</tr>
<tr>
<td></td>
<td></td>
<td>L. cerebellum</td>
<td>-26, -84, -52</td>
<td>3.16</td>
</tr>
</tbody>
</table>

The coordinates (x, y, z) locate the maxima within an area of rCBF increase associated with a given contrast; z values correspond to P < 0.01 (uncorrected for multiple comparisons). The anatomical locations and the Brodmann area (BA) estimates are derived from Talairach and Tournoux (29). R., right; L., left.
DSWE \( [F(1, 11) = 3.58, P = 0.08] \). Within the right ROI, AWL raised the blood flow more than DSWE \( [F(1, 11) = 13.59, P = 0.00] \) and SSWE \( [F(1, 11) = 33.56, P < 0.00] \). AWLO increased the rCBF more than SSWE \( [F(1, 11) = 5.02, P = 0.04] \), but not more than DSWE. Neither AWL with novel compared with old words nor DSWE vs. SSWE differed significantly in either ROI.

In the second ROI analysis, the PET difference activations resulting from each contrast were analyzed in 15 contiguous, circular coronal ROIs (2-mm slice thickness) aligned along the long axis of the left and right hippocampal formation covering a rostral-caudal length of 2.8 cm. The activations underlying the associative process were higher than the activations underlying novelty detection and depth of processing. None of the activations associated with novelty detection or depth of processing reached statistical significance in any of the left and right hippocampal ROIs. Fig. 4 displays difference activations between AWL and DSWE that were significant only in the anterior hippocampal formation (31, 32).

**DISCUSSION**

The whole-brain analysis of our PET data indicated that neither novelty detection nor depth of processing, but the associative process alone, yielded a significant increase in hippocampal activation. Remarkably, the pattern of task-induced differences in hippocampal blood flow (Fig. 3) was present in each single subject, indicating its robustness. These results are in line with those from our previous PET study (33) comparing associative picture learning and single picture learning, detecting novel picture associations, and retrieving old picture associations. Associative picture learning increased the hippocampal blood flow more than any other task. The use of the unsubtracted activity in the present ROI analysis clarifies why the comparison of AWL with either of the two Single Word Learning (SWL) tasks had yielded a right instead of the expected left hippocampal significance in the whole-brain analysis: SWL increased the blood flow in the left more than in the right anterior hippocampal formation, elevating the reference activations and, thus, decreasing the effect size in the left anterior hippocampal formation. Although the activations during single-word learning were higher in the left than in the right anterior hippocampal formation, the comparison between DSWE and SSWE did not reach statistical significance in either anterior hippocampal area, and SSWE activated both anterior hippocampal areas significantly less than AWL. AWL by itself activated the two hippocampi equally. When AWL was carried out with previously studied words, the left anterior hippocampal blood flow yielded the highest values.

![Fig. 2](image2.png)

**Fig. 2.** PET activations in the two contrasts, Associative Word Learning–Shallow Single Word Encoding (Upper) and Associative Word Learning–Deep Single Word Encoding (Lower). Coronal (y) SPM(\( z \)) maps (27, 28) at the indicated coordinate position superimposed on an MRI, both spatially normalized into stereotaxic space (29). R, right. SPM(\( z \)) maps are thresholded at 0.01. The strong activations in the left temporal cortex (Upper) are due to the semantic processing of words during AWL; they are subtracted out in the second comparison.

![Fig. 3](image3.png)

**Fig. 3.** ROI analysis within the right and left anterior hippocampal formation. Displayed is the mean activity averaged over 12 subjects during each of the four learning tasks. (Bars = SD.)

![Fig. 4](image4.png)

**Fig. 4.** Difference PET activations from the contrast of AWL to DSWE. Difference PET activations are averaged over 12 subjects for each of 15 contiguous, 2-mm-thick ROIs, which are aligned along the long axis of the right hippocampus (left side of abscissa = rostral hippocampus). (Bars = SD.) Numbers on the abscissa denote \( y \) (anterior–posterior) coordinates from the stereotaxic atlas of Talairach and Tournoux (29). The asterisk indicates significance (\( P < 0.05 \), two-tailed) of the rCBF difference within a ROI.
Interestingly, Wagner et al. (34) reported significantly greater activation during deep (deciding whether a word is abstract or concrete) than shallow (deciding whether a word is printed in upper- or lowercase letters) single-word processing in the left frontal, parahippocampal, and fusiform cortices. This comparison is similar to ours, and the left frontal activations underlying semantic processing are overlapping in the two studies. Yet, we found no significant activation differences in the parahippocampal and the fusiform cortices. This difference may be due to the unequal processing demands of the two shallow-encoding tasks; while counting vowels across words (our shallow-processing task) is time-consuming (mean = 3.975 ms) because it requires a sequential analysis of each letter string, a brief look is sufficient to determine whether a word is printed in upper- or lowercase letters (mean = 539 ms). Because our shallow-encoding task is requiring an accurate visual/sublexical analysis, there will be more concurrent fusiform and possibly parahippocampal activations during this task than during the shallow-encoding task of Wagner et al. (34). Activations in these areas underlying deep processing therefore might get subtracted out in our comparison of deep with shallow encoding.

The present results replicate and extend those of previous PET studies (33, 35–36) by demonstrating (i) that the process of semantically associating items significantly activates the hippocampal formation, (ii) that associating originally unrelated items challenges the hippocampal formation more than deep, single-item encoding and novelty detection, and (iii) that this effect is independent of the kind of stimulus material used. These results substantiate the “association” hypothesis (14–17, 24) in the human and correspond to neuropsychological findings, indicating that patients with hippocampal damage are most impaired in spatial (13) and nonspatial associative learning (37). The reason why patients with amnesia induced by hippocampal damage still fail in single-item learning is probably because single items need some associative processing as well to be recalled successfully later, e.g., they need to be associated with the learning context. In our experiment, we had varied the degree of associative processing across learning tasks and found that it was this component that drives the anterior hippocampal formation most. This result confirms the hypothesis raised by Schacter and Wagner (32) that encoding activations tend to be located in the anterior medial temporal lobe when encoding tasks require relational processing of multiple stimuli.

The finding that the hippocampal formation functions as an associator of information processed in distributed cortical regions almost can be deduced from the anatomy of the primate hippocampal system (3, 4, 14–24): the hippocampal system is hierarchically organized with the hippocampus located at the top. Inputs from the various sensory association cortices are channeled through the perirhinal and parahippocampal cortices to the entorhinal cortex (or directly to the entorhinal cortex) and, from there, to the hippocampal formation. The hippocampus is reciprocally connected with these sensory areas and therefore can reactivate them. This suggests that the hippocampus performs the most complex mnemonic computations such as associating the different inputs with one another and reactivating the involved cortical sites for consolidation and retrieval of information. In our experiment, the establishment of a semantic relationship between two primarily unrelated abstract nouns activated the hippocampus most, presumably because the hippocampus associated activations in disparate cortical regions corresponding to representations of the semantic lexicon.

Hippocampal activations in learning experiments typically were reported when targets were unfamiliar, complex, multifaceted, and nonverbal (33, 38–40). A reason for this might be that the memorization of such material strongly requires the binding of subcomponents into a complete mental represen-

tation. Most experiments using verbal stimuli have failed to demonstrate hippocampal activations (41–46). The verbal learning tasks used might not have challenged the critical function of the hippocampal formation enough. Some recent experiments on verbal learning and memory, however, yielded hippocampal activations by use of instructions that encouraged the association of a presented word with a self-created semantic context or stored concepts. In light of our findings, these experiments can be reinterpreted in terms of associative processing of information. Schacter et al. (47) reported right hippocampal activation during the recall of words that had been encoded by counting the number of meanings associated with each—this task binds the target word to stored concepts. The hippocampal activation during retrieval might reflect the recovery of these associations. Similarly, Rugg et al. (10) found left hippocampal activation during the retrieval of words that had been encoded by generating a sentence that includes the target word. This encoding task binds the presented word into the self-created sentence context. Dolan and Fletcher (48) reported left hippocampal activation during an associative learning task with word pairs that had not been studied during two preceding paired-associate learning runs. They found less left hippocampal activation if one word of each pair had been studied during both preceding runs and even less hippocampal activation if the presented word pairs had been studied twice before. Thus, hippocampal activation seemed to have increased as a function of associative learning.

In conclusion, hippocampal activations found in our and other functional imaging studies of memory can be reinterpreted as correlates of establishing (or recovering) semantic associations between components of a learning event in memory.

We thank G. K. von Schultheiss for the PET infrastructure, T. Berthold and K. Schwedler for help with the data collection, and the referees for their careful reviews. This research was supported by Swiss National Science Foundation Grant 31-47203.96.