Hippocampal Involvement in the Acquisition of Relational Associations, but Not in the Expression of a Transitive Inference Task in Mice

Ann Van der Jeugd, Hannelore Goddyn, Annelies Laeremans, Lutgarde Arckens, and Rudi D'Hooge
University of Leuven

The hippocampus (HC) has been suggested to play a role in transitive inference (TI) on an ordered sequence of stimuli. However, it has remained unclear whether HC is involved in the expression of TI, or rather contributes to TI through its role in the acquisition of the sequence of elements (Frank, Rudy, & O'Reilly, 2003). Presently, the authors compared the effects of excitotoxic dorsal HC lesions in C57BL mice that received surgery before or after they were trained to discriminate between pairs of visual stimuli. Performance on a subsequent TI task was worse in mice with pretraining lesions than in those with postraining lesions, which showed similar performance to shams without lesions. This indicates that HC is not involved in the expression of TI, but may merely help to acquire the underlying representations required for TI.

Keywords: transitive inference, hippocampus, excitotoxic lesions

Ever since the report of memory impairment in patient H.M. (Scoville & Milner, 1957), the role of the hippocampus (HC) in episodic memory has been undisputed. The HC was later found to be critical in spatial learning (O'Keefe & Nadel, 1978; Morris, Garrud, Rawlins & O'Keefe, 1982), and recent work has involved HC in various cognitive operations, although the precise definition of its involvement is still vigorously debated (Martin, de Hoz, & Morris, 2005). Transitive inference (TI) is one of these HC-dependent operations (Dusek & Eichenbaum, 1997; Van Opstal, Verguts, Orban, & Fias, 2008), which consists of deriving a conclusion from two related premise pairs that invoke a transitive relationship. In a typical rodent TI task, a sequence (e.g., A > B > C > D > E) is learned by presenting consecutive premise pairs (e.g., AB), and rewarding only the one with the highest rank in the sequence (i.e., A). When animals are subsequently presented with an untrained stimulus pair (e.g., BD), TI is evident from a choice that corresponds to the learned sequence (i.e., B).

Theorists have suggested that HC is critical for the acquisition of conjunctions of events (Meeter & Murre, 2005; O'Reilly & Rudy, 2001; Rolls & Treves, 1998). This integrates the role of HC in episodic and spatial memory as both an episode and a spatial location consist of a combination of different features that must be stored together. With respect to its role in TI, Frank, Rudy, and O'Reilly (2003) proposed a computational model in which, during the training phase of a TI task, HC would control the acquisition of elementary conjunctions. These conjunctions are used to build up a gradient of associative strengths in neocortex, which can be used subsequently for solving transitive pairs without further HC involvement.

The present study uses a visual discrimination task in C57BL mice to test this hypothesis, implementing a design already proposed by Frank et al. (2003). Dusek and Eichenbaum (1997) used olfactory stimuli to train rats on a set of ordered premises. Performance on transitive pairs was compared between rats that received HC lesions before training and rats with sham lesions. Despite similar performance on premise pairs, rats with pretraining lesions performed worse on the TI pairs than animals with sham lesions. However, it cannot be excluded that the observed TI impairment was due to an acquisition defect during the training phase. We, therefore, included one critical group of animals with postraining lesions. If HC is exclusively involved in the acquisition of element pair conjunctions, mice with pretraining lesions should display deficits during the TI test, whereas mice with postraining lesions should perform at the level of controls. Conversely, if HC is also essential for the expression of TI, the groups with lesions pre- or postraining should both perform worse than controls.

Method

We used 27 female C57BL/6J mice (Elevage Janvier, Le Genest Saint Isle, France). Animals were housed in standard plastic cages...
under standard lab conditions with a 12:12 light/dark cycle and ad libitum access to water. Mice from the pre- and posttraining conditions were trained and tested together and randomly distributed over two of the cages; the experimenter was condition-blind. During shaping, training, and testing, the animals were reduced to 80–90% of their original body weight. All experiments were performed between 9 and 16h during the light phase of their activity cycle. The experiment was approved by the Ethical Committee of the K. U. Leuven. Five animals died during the experiment due to epileptic seizures (two from the pretraining lesion condition, three from the posttraining lesion condition). As a result, the number of animals was 8, 7, and 7, in pretraining-lesion, posttraining-lesion, and sham conditions, respectively.

**Surgery**

We only included lesions in HC because Dusek and Eichenbaum (1997) showed that fornix or perirhinal/entorhinal lesions led to equivalent deficits in Ti performance. A first intracerebral injection was performed before training (3 days after the shaping phase). Mice of the pretraining condition received bilateral HC lesions, whereas mice of the posttraining and sham conditions received sham treatment. A second injection was performed immediately after training. Now, the pretraining and sham mice obtained sham treatment, whereas the posttraining-lesion group obtained bilateral HC lesions.

Surgical procedures were as described before (Goddyn, Leo, Meert, & D’Hooge, 2006). Briefly, the animals were anesthetized by intraperitoneal injection of xylazin (10 mg/kg) and ketamin (100 mg/kg) in physiological saline. After sagittal incision of the scalp, animals were placed in a stereotactic apparatus (Narishige Scientific Instruments, Tokyo). A craniotomy was made 2.5 mm posterior to bregma and 2 mm lateral to the midline, and 75 nmol quinolinic acid (in 0.5 µl phosphate-buffered saline) was injected at a depth of 1.8 mm with a 1-µl syringe (Hamilton, Reno, NV). After the injection, the skin was disinfected and the scalp sutured. Two days later the contralateral HC was injected. The sham procedure was similar in all respects, except that no quinolinic acid was injected.

**Apparatus and Protocol**

Mice were trained in a visual discrimination task using a T maze with arms of 50 cm long and 5 cm wide. Each arm was closed by sliding doors that were located at 10 cm from the junction. Graphical images (5 × 5 cm²) were fixed to the doors of the lateral arms with each of the five images consisting of 50% white and 50% black surface (see Figure 1). A different ordinal position of the stimuli (from A to E) was used in each cage. During the shaping (before surgery), mice were placed in the central arm with their noses facing the junction and the first sliding door was opened after 3 s. The animals were initially trained to run into the lateral arms of the T maze using a small chocolate cornflake as a reward. Then they were trained to choose the arm that was closed by a white door and avoid the arm with a gray door. A choice response was only registered when the door was touched with their nose and the door was lifted to provide access to the reward. Each mouse was allowed only one choice per trial, after which the mouse was returned to its cage, and the maze was thoroughly cleaned for the next mouse. Each trial lasted a minute with five minutes intertrial interval. The correct side was chosen randomly on each trial. Each mouse performed 20 trials per day, and shaping ended when they reached a criterion of 80% correct.

Training started one week after the first operation. Two stimuli were shown per trial from which the animal was to choose one. Learning the premises consisted of four parts. In part 1, there were five presentations of pair AB, followed by five presentations of pair BC, and so on. This was repeated until each animal had reached a score of 15/20 (75%). Part 2 was similar except that only three consecutive presentations of each pair were given. The criterion to go to the next part was 9/12 (75%). In part 3, only one presentation of each pair was given in a row; the criterion to go to the next part was 9/12 correct. Finally, in part 4, pairs were presented randomly in sequences of 16 stimuli. This part ended when the animals reached the criterion of 12/16 (75%) correct.

One week after the second operation, the test phase started. During the test phase, five sequences, one on each day, were presented to each mouse. Each sequence contained 20 trials, consisting of four training pairs each shown four times per sequence, and two untrained probe pairs (BD and AE), each shown twice per sequence. Hence, the critical transitive pair BD was presented on average once every 10 trials. All pairs were randomly intermixed. Stimuli were rewarded consistent with their linear order (i.e., B and A, respectively). Five such sequences were presented to each mouse.

**Lesion Site Confirmation**

To check whether the injections had a behavioral effect, we measured cage activity before shaping and after each operation. All of the mice with assumed lesions showed the hyperactivity typical for HC damage (Goddyn et al., 2006). After the experiment, the animals were killed with CO₂ and decapitated. Brains were removed and postfixed in 4% formaldehyde solution. Coronal sections (50 µm thick) were cut on a vibratome, placed on gelatin-coated slides, air-dried and Nissl-stained with 1% cresyl violet (Fluka Chemical, Sigma-Aldrich) according to standard protocols. They were then dehydrated, cleared and coverslipped. These sections were examined under a Leitz DM RBE microscope (Leica, Leitz Instruments, Heidelberg, Germany) for histological verification of the lesions.

**Results**

Shaping took part in nine days; the number of trials needed for shaping (122.5, 120, and 122.9 for pretraining, posttraining, and
ACQUISITION OF RELATIONAL ASSOCIATIONS

sham groups, respectively) did not differ significantly between the three groups (F < 1). On the ninth day of shaping, animals from the pretraining, postraining and sham groups showed 87.5%, 88.3%, and 83.3% accuracy, respectively, and did not differ significantly (F_{2, 21} = 1.488, p = .251).

We next performed an ANOVA on the number of trials to reach criterion (TTC) in the training phase. We followed the logic of Frank, O’Reilly and Curran (2006) and included the factors group (pre/post/sham), part (1 vs. 2–4), and pair type (inner pairs BC & CD vs. outer pairs AB & DE). Results are shown in Figures 2 (collapsed over pair type) and 3 (collapsed over part). Referring to Figure 2, the group variable was significant (F_{2, 19} = 3.899, p < .05): Mice in the pretraining lesion group were slower to reach criterion than the other two groups. Focused statistics confirmed that there was a significant difference in TTC performance between mice with lesions pre- versus posttraining (one-tailed t_{13} = 2.459, p < .05), but not between mice with postraining lesions versus sham lesions (one-tailed t_{13} = 2.459, p < .05), although this is difficult to interpret given that the criterion was different in the different parts (see Methods). However, there was also a marginally significant part \times group interaction (F_{2, 19} = 3.222, p = .062): Group effects tended to be larger in part 1 (e.g., pretraining—sham = 42.710 TTC, t_{13} = 2.400, p < .05) than in part 2 (e.g., pretraining—sham = 0.760 TTC, t_{13} = .188, p = .854). There was also an effect of pair (see Figure 3, F_{1, 19} = 21.484, p < .001): Acquiring inner pairs (90.9 TTC) took longer than outer pairs (74.1 TTC). There were also interactions between pair and group (F_{2, 19} = 8.623, p < .01) and between pair and part (F_{1, 19} = 21.705, p < .001). Finally, there was a three-way interaction between the three factors (F_{2, 19} = 5.917, p < .05).

In view of the role of HC in rapidly encoding stimulus conjunctions in very few trials (O’Reilly & Rudy, 2001), and the significant interaction between pair, group and part, we further investigated the group and pair effects in part 1 of the training (see Figure 4). In this early stage of learning, all three groups need an approximately equal number of TTC for learning the outer pairs (F < 1). However, animals with pretraining lesions were impaired at acquiring the inner pairs, compared to animals without HC lesions at that stage, or shams (F_{2, 19} = 6.385, p < .05), similar to what was found by Frank et al. (2006) in humans. Additional analyses verified that there was a significant difference in TTC performance on inner pair items between mice with pretraining lesions versus postraining lesions (one-tailed t_{13} = 2.687, p < .05), and between mice with pretraining lesions versus sham lesions (one-tailed t_{13} = 4.87, p < .001), but not between mice with postraining lesions versus sham lesions (t_{12} = 1.41, p =

![Figure 2](image-url)  
**Figure 2.** Mean number of trials required to reach the criterion during premise training for each group (pre = pretraining lesion condition; post = postraining lesion condition; sham = sham condition). Block sizes of five, three, one, and random correspond to training parts 1, 2, 3, and 4, respectively (see text for full explanation). Errors bar denote one standard error of measurement (SEM).

![Figure 3](image-url)  
**Figure 3.** Mean (±SEM) number of trials required to reach the criterion for the outer premise pairs (AB and DE) and the inner premise pairs (BC and CD) for each group.

![Figure 4](image-url)  
**Figure 4.** Mean (+SEM) number of trials required to reach the criterion for the outer premise pairs (AB and DE) and the inner premise pairs (BC and CD) during the first part of training for each group.
In the part 2 data, all focused t-statistics were nonsignificant at the .05 level. In Figure 5, we plot the percentage of correct responses on the premise pairs during the test phase. There was a significant effect of pair ($F_{3, 57} = 3.882, p = .014$), but not of group ($F_{2, 19} = .399, p = .676$), and there was no interaction between the two factors ($F_{6, 114} = .540, p = .775$).

Performance on the test pairs (BD and AE) is shown in Figure 6. Mice with posttraining lesions as well as sham-treated mice performed above chance level on the critical BD pairs ($t_b = 3.667, p = .010; t_s = 3.122, p = .021$, respectively), indicating that mice in these two treatment groups were able to use TI to solve these problems. Notably, this was not the case in the mice with pretraining lesions ($t_7 = .753, p = .704$). Focused statistics confirmed that there was a significant difference in BD performance between mice with pre- versus posttraining lesions (one-tailed $t_{11} = 1.79, p = .048$), and between mice with pretraining lesions and sham-treated mice (one-tailed $t_{13} = 1.926, p = .038$), but not between mice with posttraining lesions and sham-treated mice ($t_{13} = .390, p = .704$). On the AE test pairs, all three groups performed significantly different from chance level ($p < .001$; Figure 6). However, AE pair performance is not considered further, because performance on AE pairs can be attributed to reinforcement history of the individual stimuli (A was always rewarded, E was never rewarded), transitive inference, or a combination of both. In Figure 6, we also depict the performance on premise pairs BC and CD. Performance on these pairs was comparable to that on BD pairs in the posttraining lesion group ($p > .8$), and in the sham lesion group ($p > .8$), but not in the pretraining lesion group ($t = 2.4; p < .05$).

Histological evaluation of HC damage in mice with pre- versus posttraining lesions indicated that the injections induced bilateral damage that was largely confined to CA1 region (see Figure 7). Microscopic visual inspection of all sections revealed that between 30 and 100% of CA1 pyramidal cell layer was destroyed. In 40% of the mice, CA1 was completely destroyed, whereas dentate gyrus granule cells were spared in most animals throughout anteroposterior HC. In some mice, minor unilateral cell loss was observed in dentate gyrus and CA2/3 at bregma −2.5 mm, the site of quinolinic acid injection. This can be attributed to the fact that, at this level, CA2/3 and dentate gyrus are in very close proximity to CA1.

**Discussion**

We have shown that animals with damage to HC before training perform worse on TI problems than animals with damage after training. In fact, the latter animals performed similar to sham-treated mice. These results support the view that HC is not involved in the expression of TI (i.e., making inferential judgments), which would have been consistent with a difference between posttraining and sham lesions. We also observed that the deficit of the pretraining lesion animals in the training phase was mainly located in the inner pairs (BC, CD) rather than in the outer pairs (AB, DE), and specifically in the early phase of training. This is again consistent with a role for HC in acquiring the appropriate stimulus representations. Indeed, the outer pairs but not the inner pairs can be solved by a simple reinforcement-based rule (i.e., A was always reinforced, E never).

It is generally assumed that HC is part of an evolutionary conserved neural network for complex associative learning (Manns & Eichenbaum, 2006). In line with this, functional imaging in humans recently showed that HC activity during the initial stage of working memory maintenance contributed to long-term memory formation (Ranganath, Cohen, & Brozinsky, 2005). Many authors have indicated that HC controls working memory functions and the encoding of information (Wan, Pang, & Olton, 1994; Karlsgodt, Shirinyan, van Erp, Cohen, & Cannon, 2005), particularly the rapid storage of conjunctions of features (O’Reilly & Rudy, 2001; Rolls & Treves, 1998; for evidence at the single-cell level, see Ekstrom et al., 2003). Some authors found HC activation during a TI task in human subjects (e.g., Greene, Gross, Elsinger, & Rao, 2006; Heckers, Zalesak, Weiss, Ditman, & Titone, 2004; Van Opstal et al, 2008), but the nature of HC involvement cannot be precisely defined in these studies. For example, Heckers et al. (2004) found that infering...
A > C from A > B and B > C led to more HC activation than inferring A > D from A > B and C > D (in which case the two sequences A > B and C > D were required overlap). However, the latter task can be solved by just knowing that A is leftmost in the sequence, and hence, HC activation may merely relate to this. Another imaging study found medial temporal lobe activation while participants were learning (a subset of) the adjacent stimulus pairs, but not during TI itself (Nagode & Pardo, 2002).

Also other studies indicated that the role of HC in transfer tasks might be comparable to that presently proposed for TI tasks. Such transfer tasks could be solved with the use of associative values of the elemental stimuli that were acquired during initial training (Sutherland & Rudy, 1989; O’Reilly & Rudy, 2001). Driscoll, Sutherland, Prusky, and Rudy, (2004) trained rats on a visual discrimination task, and subsequently presented them with novel combinations of the stimuli used during acquisition. They found that rats with HC damage before training were perfectly able to solve these transfer tasks, contrary to what Eichenbaum, Mathews, and Cohen (1989) demonstrated in rats with HC lesions using an odor discrimination task. Some authors argued that the discrepancies between the two studies can be attributed to differences in the approach of animals with HC lesions and control animals in the odor discrimination task (Driscoll et al., 2004; Rudy & Sutherland, 1994). Thus, the difference observed by Eichenbaum et al. (1989) in the transfer task could be attributed to differences in acquisition of the discrimination task, rather than to the involvement of HC in transfer tasks.

However, it is less obvious how TI can be understood in a conjunctive coding framework. HC might help to store the different elements with their subsequent links, and TI could be based on HC-dependent “walking” through the sequence. This would mean that pairs become more difficult with increasing distance between the elements, but this prediction is contrary to what is experimentally observed (Van Elzakker, O’Reilly, & Rudy, 2003). In a model by Frank et al. (2003), HC is critical for learning the premise pairs, but not the expression of TI. They propose that a gradient of associative strengths is constructed in neocortex, which is required for solving TI pairs. Similar proposals appear both in the animal learning literature (e.g., von Fersen, Wynne, Delius, & Staddon, 1991) and in the human numerical cognition literature (Verguts, Fias, & Stevens, 2005). In the present context, the gradient that develops without HC is less steep than that which develops in the intact brain, causing mice with HC lesions to fail on TI test pairs, even when they succeed on the training pairs. The present study is the first to validate this model empirically.

It has to be noted that the present findings could be characteristic for lesions of the dorsal CA1 region. Excitotoxic HC lesions typically damage dorsal HC (mainly CA1), while sparing other parts of the HC formation (Deacon, Bannerman, Kirby, Croucher, & Rawlins, 2002). Recent lesion studies demonstrated functional specialization in different subregions of dorsal HC. For example, rats with lesions to either dorsal CA1 or CA3 were tested in a task that interposed a 10-s delay between the presentation of an object and an odor (Kesner, Hunsaker, & Gilbert, 2005). Rats with dorsal CA3 lesions learned this object-trace-odor task as promptly as controls, whereas rats with lesions to dorsal CA1 were impaired. Such evidence further confirms that CA1 is directly involved in chunking information across time, generating units based on a specific order of occurrence.

In conclusion, HC can be thought of as an associative device that integrates inputs from different sensory modalities, and provides temporal and spatial contexts for these associations in tasks that require such an integration. However, the present data and similar work suggest that its involvement is restricted to the acquisition of the relevant information (e.g., premise pairs) required for accomplishing such tasks.

References
Driscoll, I., Sutherland, R. J., Prusky, G. T., & Rudy, J. W. (2004). Damage to the hippocampal formation does not disrupt representational flexibility as measured by a novelty transfer test. Behavioral Neuroscience, 118, 1427–1432.


Received April 30, 2008
Revision received August 19, 2008
Accepted August 25, 2008