



## UvA-DARE (Digital Academic Repository)

### The ventral striatum in goal-directed behavior and sleep: intrinsic network dynamics, motivational information and relation with the hippocampus

Lansink, C.S.

**Publication date**  
2008

[Link to publication](#)

#### **Citation for published version (APA):**

Lansink, C. S. (2008). *The ventral striatum in goal-directed behavior and sleep: intrinsic network dynamics, motivational information and relation with the hippocampus*.

#### **General rights**

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

#### **Disclaimer/Complaints regulations**

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

## CHAPTER 4

### Preferential reactivation of motivationally relevant information in the ventral striatum

Carien S Lansink, Pieter M Goltstein, Jan V Lankelma, Ruud NJMA Joosten, Bruce L McNaughton and Cyriel MA Pennartz.

*Journal of Neuroscience* (2008) 28(25):6372-6382



**Abstract**

Spontaneous 'off-line' reactivation of neuronal activity patterns may contribute to the consolidation of memory traces. The ventral striatum exhibits reactivation and has been implicated in the processing of motivational information. It is unknown, however, whether reactivating neuronal ensembles specifically recapitulate information relating to rewards that were encountered during wakefulness. We demonstrate a prolonged reactivation in rat ventral striatum during quiet wakefulness and slow-wave but not REM sleep. Reactivation of reward-related information processed in this structure was particularly prominent and this was primarily attributable to spike trains temporally linked to reward sites. It was accounted for by small, strongly correlated subgroups in recorded cell assemblies and can thus be characterized as a sparse phenomenon. Our results indicate that reactivated memory traces may not only comprise feature- and context-specific information but also contain a value component.

**Introduction**

82 The Ventral Striatum (VS) is one of the key structures involved in the motivational control of behavior (Mogenson et al., 1980; Robbins and Everitt, 1996). Evidence indicates that the VS is required for the learning of cue-outcome (Parkinson et al., 1999b) and response-outcome (Kelley et al., 1997) associations to invigorate or guide goal directed behavior on the basis of the motivational value of cues and contexts. Single-unit recording studies in awake rodents and primates have established VS neural responsivity to actual and expected rewards or aversive reinforcers (Schultz et al., 1992; Roitman et al., 2005), cues predicting these reinforcers (Tremblay et al., 1998; Setlow et al., 2003; Roitman et al., 2005) and motor responses required to obtain or avoid these (Shidara et al., 1998; Tremblay et al., 1998; Hassani et al., 2001). Recent fMRI results have confirmed the role of the VS in processing and predicting reinforcers in a time-specific manner (O'Doherty et al., 2006).

Recently, the VS was indicated to participate in memory consolidation. Pharmacological interventions in protein synthesis, glutamatergic or dopaminergic neurotransmission in the VS shortly after training impaired instrumental (Hernandez et al., 2002), spatial (Setlow and McGaugh, 1998; Sargolini et al., 2003) and Pavlovian approach learning (Dalley et al., 2005). In addition, spontaneous reactivation of neuronal firing patterns (i.e. in the absence of external stimuli) occurs in the VS during sleeping periods following a behavioral experience (Pennartz et al., 2004). In these 'off-line' periods, the recurrence of neuronal activity patterns might contribute to memory consolidation by strengthening synaptic connections activated during the preceding behavior or by forming more direct

connections among items stored in distributed form throughout the brain (Pavlidis and Winson, 1989; Wilson and McNaughton, 1994; McNaughton, 1998; Rasch et al., 2007).

Ventral striatal reactivation, similar to hippocampal replay (Kudrimoti et al., 1999), is manifested during periods of slow wave sleep and occurs especially in neuronal subgroups whose firing rates are modulated in close temporal association with sharp wave-ripple complexes, i.e. high-frequency oscillations in the hippocampal Local Field Potential (LFP; O'Keefe and Nadel, 1978; Buzsaki, 1986). It is unknown, however, whether VS reactivation preferentially takes place during ripple episodes and whether it also occurs during Rapid Eye Movement (REM) sleep. Interestingly, although a significant reactivation effect was reported, its occurrence and strength appeared variable across sessions (Pennartz et al., 2004).

Thus far, replay has been primarily studied in populations of hippocampal pyramidal cells exhibiting 'place fields' (Wilson and McNaughton, 1994; Skaggs and McNaughton, 1996), where it may be assumed to pertain to spatial and contextual information processing. For the VS however, the behavioral correlates of the reactivated information are unknown. We hypothesized that it is reward-related information that is reprocessed 'off-line' in this area to endow the memory trace with a motivational component. We tested this by examining whether the VS specifically reactivates reward-related information rather than the overall ensemble spike patterns that occur during behavior in general. In addition, we examined which sleep stages exhibit this reactivation, including periods of REM.

83

## **Materials and Methods**

### ***Experimental procedures***

*Behavior* Four male Wistar rats (375-425g, Harlan, the Netherlands) were individually housed under a 12/12h alternating light-dark cycle with light onset at 8:00 AM. All experiments were conducted in the animal's inactive period. On training and recording days, intake of water was limited to a two hour period following training or recording. Food was available *ad libitum*. Before surgery, rats were pre-trained on a linear track (185cm long x 10cm wide, 40 cm elevated from the floor) to shuttle back and forth for reinforcements available at both ends of the track. Over the course of training rats were introduced to three kinds of rewards which differed in both taste and texture (sucrose solution 10%, vanilla desert and chocolate mousse) and to a partial reinforcement schedule.

During recordings, rats were subjected in daily sessions to a protocol consisting of a rest period (pre-behavioral rest, 20-60 min) followed by a phase of reward searching

behavior on a triangular track (Track, 20 min) and concluded by a second period of rest (post-behavioral rest, 60-120 min). The track (equilateral sides 90 cm, width 10 cm) was novel to the rats at the first recording session. On the track, the rats were required to run in one direction, stopping only at reward wells positioned in the center of each arm to check whether a reward was available. To promote differential firing associated with reward sites, the three types of rewards used during pre-training were also provided to the animal in the task. Each lap, one of the three reward types could be obtained from its corresponding well. The combination of reward type and well location was fixed throughout all sessions. Rats spent the rest episodes on a towel folded in a wide flowerpot situated next to the track.

*Surgery and recordings* Rats were implanted with a multi-electrode microdrive containing 7 individually movable tetrodes directed to the VS (1.8 mm anterior and 1.4 mm lateral to Bregma (Paxinos and Watson, 1986) whereas additional tetrodes were placed in the hippocampus. Reference electrodes were placed in the corpus callosum, and near the hippocampal fissure. A skull screw located on the caudal part of the parietal skull bone contralateral to the drive location served as ground. Spike trains from individual cells, LFPs and the position of the rat were recorded using a 64 channel Cheetah recording system (Neuralynx, Bozeman, MT, USA). When signals exceeded a manually preset voltage threshold, waveforms were sampled at 32 kHz for 1 ms (filter settings: 600-6000Hz). LFPs were continuously sampled at 1690 Hz and band-pass filtered between 1 and 475 Hz. Using an array of light-emitting diodes on the headstage, a video tracking system extracted the rat's position on the maze at 60 Hz with a resolution of 2.5 mm/pixel. The behavior of the rat was also stored on videotape. All experimental procedures were in accordance with national guidelines on animal experimentation.

84

### ***Data Analysis***

*Spike sorting* Spikes from neurons were separated from those emitted by other neurons recorded on the same tetrode by grouping spikes with similar distributions of waveform properties across the four channels of a tetrode using standard automated and manual clustering methods, i.e. BubbleClust and MClust respectively. BubbleClust groups spikes based on nearest-neighbor distances, clustering spikes that are close to each other, given features of the waveform such as peak-amplitude or area under the curve and principal components of a spike on each tetrode channel. MClust facilitates manual selection of clusters by allowing users to limit cluster membership based on boundaries drawn on two-dimensional plots of the waveform features. Clusters of spikes were attributed to a single unit on the basis of waveform characteristics and when they exhibited less than 0.1% of spike-intervals within a 2 ms refractory period in their inter-spike interval histograms (Fig

4.1A). Units were only included in the analyses when they emitted at least 20 spikes in each behavioral/rest episode. Putative interneurons were distinguished from principal cells by means of average firing rate ( $>8$  Hz) and waveform characteristics (small peak-to-valley width, valley shape) and were not included in analysis.

*Identification of rest and sleep phases / ripple detection* Pre- and post-behavioral rest phases comprised all periods of motionless behavior of the rats when they were situated in the flowerpot. Time frames with body movements during these episodes indicated by video tracker equipment and video tapes were extracted from the pre- and post-behavioral recordings. Within these rest episodes of behavioral immobility, sleep phases were identified using LFP traces recorded near the hippocampal fissure and pyramidal layer. REM sleep was primarily defined by an elevated ratio ( $> 0.4$ ) of spectral power density in the theta band (6-10 Hz) to the overall power; borders of theta oscillations were refined upon secondary visual inspection. Periods of SWS were identified by the presence of Large Irregular Activity and ripples in the hippocampal LFP (Vanderwolf, 1969; O'Keefe and Nadel, 1978; Buzsaki, 1986; e.g. Kudrimoti et al., 1999; Pennartz et al., 2004). After filtering a LFP trace from the pyramidal cell layer between 100 and 300 Hz, a ripple was detected each time the squared LFP trace exceeded a preset threshold (3.5 SD) for at least 25 ms. Adjacent ripples were merged when the ripple-interval was less than 100 ms. The presence of large irregular activity in the LFP was verified upon off-line visual inspection. Short periods of quiet wakefulness may have been included in SWS episodes because the LFP patterns of both states share principal features. For this reason, this state is further referred to as quiet wakefulness – slow wave sleep (QW-SWS). Rest periods lasting shorter than 20 s were excluded from classification and further analysis. Thus, rest periods contained QW-SWS and REM sleep as main components, whereas segments of unclassified rest period constituted a very minor part of rest. Periods of active wakefulness within the rest episodes were not analyzed as a separate state of the sleep-wake cycle because the periods were generally short and did not contain sufficient spike counts.

85

*Reward-related firing patterns* To identify reward-related units, peri-event time histograms were constructed for the rewarded and non-rewarded condition for each reward site. The histograms were synchronized on reward site arrivals, which were signaled by the crossings of off-line installed 'virtual photobeams', positioned right before the point at which the rat reached each reward well. Reward-related responses were assessed within a period of 1s before and 1s after arrival at a reward site. Spike counts were binned in 250 ms intervals. The eight bins comprising the reward period were each compared to three bins taken from the corner passage opposite to the well under scrutiny within the

same lap (Wilcoxon's matched-pairs signed rank test,  $p < 0.01$ ). A bin of the reward period was only considered significantly different when the rank test (which includes as entries a list of all spike-count values from the test bin paired with those from the control bin per reward period) indicated significance from each of the three control bins. We verified that the firing in the control period of three bins was not marked by specific deviations from the firing in all intermediate segments between corners and reward sites using peri-event time histograms and plots of the spatial distribution of firing rates. Responses were qualified as significant when one or more bins in the reward period were significantly different from each of the three reference bins. This control period was preferred over for example the average firing rate per lap because many neurons were virtually silent during track running except for their brief, phasic response at one or more reward sites. Thus, the average firing rate of these cells strongly depends on the response intensity itself, which would enhance the bias towards false negative responses (i.e. erroneously identified as non-responsive) if it were used as control value. However, results were comparable when other control measures were used such as the baseline firing rate or the average firing rate per lap. Differences between responses at the three reward sites were statistically evaluated with a Kruskal-Wallis test ( $p < 0.05$ ) followed a Mann-Whitney's U-test (MWU:  $p < 0.05$ ), whereas rewarded versus non-rewarded conditions were compared using MWU ( $p < 0.05$ ).

86

*Quantification of reactivation* The assessment of covariation in firing rates and the quantification of reactivation with the Explained Variance method was previously described (Kleinbaum et al., 1998; Kudrimoti et al., 1999; Pennartz et al., 2004; Tatsuno et al., 2006). Briefly, spike trains of simultaneously recorded neurons were binned in intervals of 50 ms to obtain sequences of spike counts for each episode. Temporal correlations of the firing patterns of neuron pairs were determined by computing Pearson's correlation coefficients for each episode separately. All coefficients of a particular rest/active episode were assembled into a single matrix and the similarity between the three matrices was determined by computing a correlation coefficient for each of three possible combinations of two rest/active episodes. These matrix-based correlation coefficients were used to determine the degree to which the variance in the correlation pattern in post behavioral rest can be explained by the pattern established during the behavioral experience while factoring out any correlations present before the behavioral experience. This quantity is expressed in the Explained Variance (EV) measure:

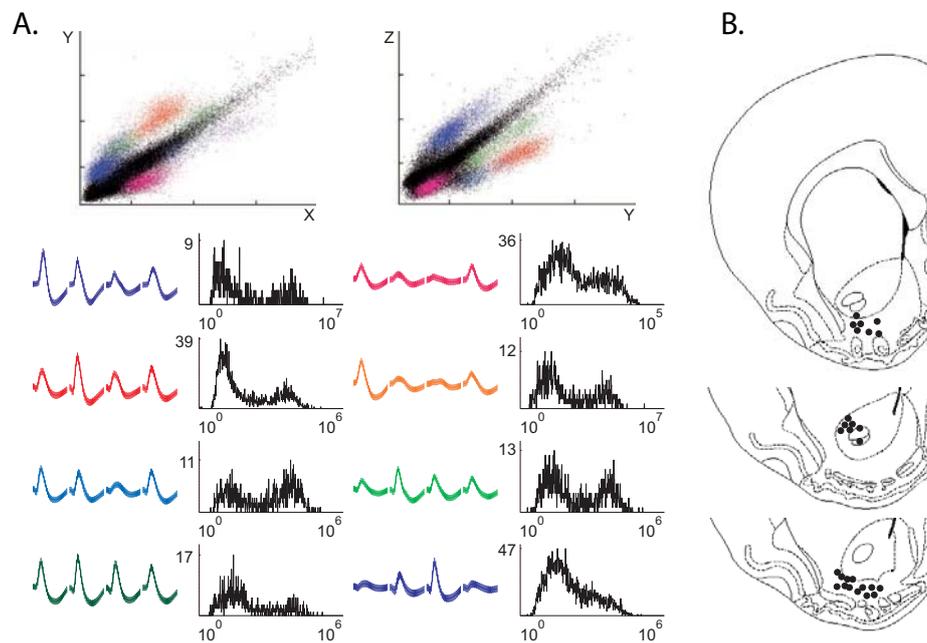
$$EV = r_{Track,R2|R1}^2 = \left( \frac{r_{Track,R2} - r_{Track,R1}r_{R2,R1}}{\sqrt{(1 - r_{Track,R1}^2)(1 - r_{R2,R1}^2)}} \right)^2 \quad (\text{EQ 4.1})$$

where R1 is the pre-behavioral rest phase and R2 is the post-behavioral rest phase. For example  $r_{\text{Track,R2}}$  equals the matrix-based correlation between the track running and post-behavioral rest pattern. EV equals the square of the partial correlation coefficient and is bounded between 0 and 1. As a within-subject and session control measure, the Reverse Explained Variance (REV) can be computed by swapping R1 and R2 in the previous equation, thereby switching the temporal order of episodes. EV and REV values were computed for all recorded sessions that contained at least 5 well-isolated active neurons and for time blocks of 20 minutes composed of quiet rest and sleep, i.e. periods of active behavior were excluded. Therefore, correlated firing between cells caused by behaviors irrelevant to the task, such as grooming, cannot influence reactivation measures. Sessions that showed reactivation ( $EV > REV$ ) in the first 20 minute rest block after track running were used to assess decay and the contribution of individual cell pairs to the session EV. Two control procedures were performed to check whether the observed reactivation was time- and cell-specific (Louie and Wilson, 2001): 1) entire spike train vectors from the behavioral episode were temporally shifted relative to the original time stamps of the same cell. The shift was circular, so that data removed from the pattern at one end were reinserted at the opposite end; the temporal distance ranged between 2s for- and backwards and varied randomly amongst cells. 2) Entire spike train vectors from the behavioral episode were randomly reassigned between cells. In both cases the temporal order of spike activity within each train was preserved. Differences between the EV and REV session values were statistically assessed with Wilcoxon's matched-pairs signed rank test.

87

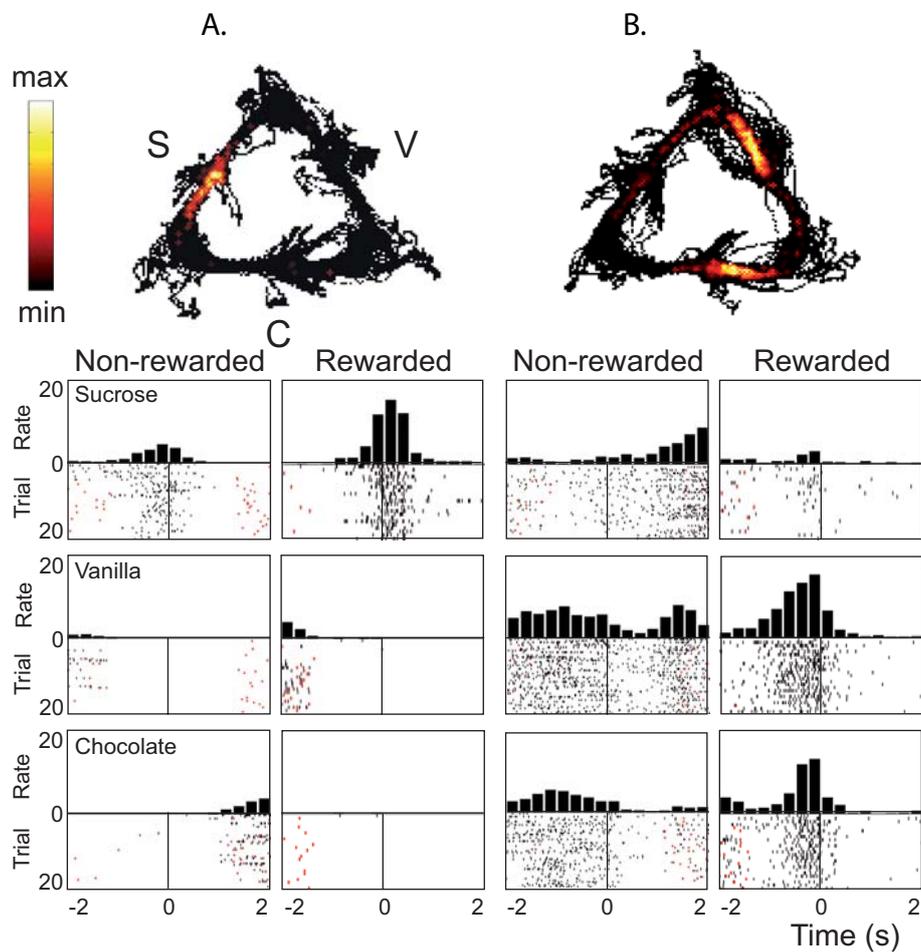
For reactivation analysis of subgroups of neurons, Pearson's correlation coefficients obtained from all rats across all sessions were pooled and subsequently EV and REV values were determined. To obtain estimates of the mean and variance of the EV and REV values, a bootstrapping procedure was applied in which randomly drawn samples were generated ( $n = 10000$ ) from the observed set of correlation coefficients (Sokal and Rohlf, 1995; cf. Hoffman and McNaughton, 2002). The re-sampling procedure was done with replacement so that each sample may contain repetitions of some triplets and omissions of others. Random samples were of the same size as the original and triplets of correlation coefficients obtained for the three task episodes (i.e. from the pre-behavioral rest, running period and post-behavioral rest) of a single recording were kept together during the re-sampling. Reactivation measures were computed for each sample resulting in distributions of estimated EV and REV values for each subset. Differences between the means of the distributions of subsets were statistically evaluated with MWU.

*Temporal order of firing* We used temporal bias (Skaggs and McNaughton, 1996) and sliding template (Louie and Wilson, 2001; Tatsuno et al., 2006) analyses to assess whether the temporal order of firing within striatal cell pairs was preserved from track running to the post-behavioral rest episode as was previously described for the hippocampus. However, probably because of the limited number of strongly reactivating cell pairs in each session and the generally low firing rates of the neurons, these analyses did not yield additional results shedding light on this question.



**Figure 4.1**

*A: Example of a tetrode recording: projection plots of cluster isolation, waveforms and inter-spike interval (ISI) histograms. Spikes belonging to separate clusters were identified according to multiple waveform features including area under the curve (plotted in this figure), peak amplitude, and principal components. X, Y and Z correspond to three out of four leads of the tetrode. The color of the waveforms shown in the first and third column from the bottom panel corresponds to the colors of the individual clusters in the projection plots. The black diagonal band corresponds to unclassified events, including noise. In the second and fourth column from the bottom panel, ISI histograms are shown for each unit represented to the left, with ISI counts on the ordinate and interval duration on the abscissa. B: Schematic representation of the endpoints of the tetrode tips. The tetrode endings were in the ventral striatum approximately between 2.2 and 1.2 anterior to Bregma and between 1.6 and 3.0 laterally as compared to an atlas of the rat brain (Paxinos and Watson, 1996).*



**Figure 4.2**

*Ventral striatal units showing firing patterns associated with one or more reward sites. Rats ran along a triangular track to obtain three types of reward (S: sucrose, V: vanilla and C: chocolate). Upper panels: spatial distribution of firing rates of two individual neurons. Local firing rates ranged from 0 to maxima of 19 Hz (A) and 18 Hz (B). Lower panels: peri-event time histograms for both cells synchronized on reward site arrivals. Rows represent different reward types whereas columns differentiate between the presence and absence of reward. Firing rate in bins of 250 ms. Red ticks indicate arrivals at other reward sites. A: Neuron showing an increased firing rate shortly before and after arriving at the sucrose reward site, mainly when a reward was obtained. B: Neuron increased its firing rate prior to arrival at two sites differentially for the rewarded versus non rewarded condition. The response at the vanilla site was stronger than at the chocolate site. The rat may have detected the availability of reward before arriving at the well by visual or olfactory cues although reward cups were filled outside the rat's field of view.*

**Table 4.1: Numbers of reward-related responses**

	Number of cells	% of total recorded cells	% of reward related responses
<i>Total number of recorded cells</i>	385	100%	
<i>Number of reward-related units</i>	66	17.1%	100%
<i>Firing rate response</i>			
Increase	61		92.4%
Decrease	5		7.6%
<i>Response timing relative to reward site arrival</i>			
Prior to arrival	20		30.3%
After arrival	20		30.3%
After arrival	26		39.3%
<i>Reward availability</i>			
Reward present	39		59.1%
Reward absent	9		13.6%
Differential responses between conditions	5		7.6%
<i>Reward sites</i>			
1 site	21		31.8%
2 sites	9		13.6%
3 sites	6		9.1%
Differential responses between sites	9 of 15		60.0%
<i>Reward availability &amp; reward sites</i>			
Differential firing to presence and sites	26		39.3%
Firing pattern not specific for any condition	8		12.1%

◀ Overview of the number of significant reward-related correlates. A total of 66 reward-related correlates were identified from 385 units (Wilcoxon's matched-pairs signed rank test,  $p < 0.01$ ). Subsets of reward-related correlates, classified under Reward availability and Reward sites, responded differentially to the presence/absence of reward or to the reward sites respectively, whereas the remaining neurons were less selective in their firing profiles. The term 'differential responses' between reward presence and reward absence indicates that in both conditions a significant response was observed but the amplitude of the responses was significantly different (Kruskal-Wallis,  $p < 0.05$ , Mann-Whitney-U test,  $p < 0.05$ ). The same specification holds for 'differential responses' between reward sites. Some neurons discriminated between reward presence/absence and between reward sites. These are listed under 'Differential firing to presence and sites'.

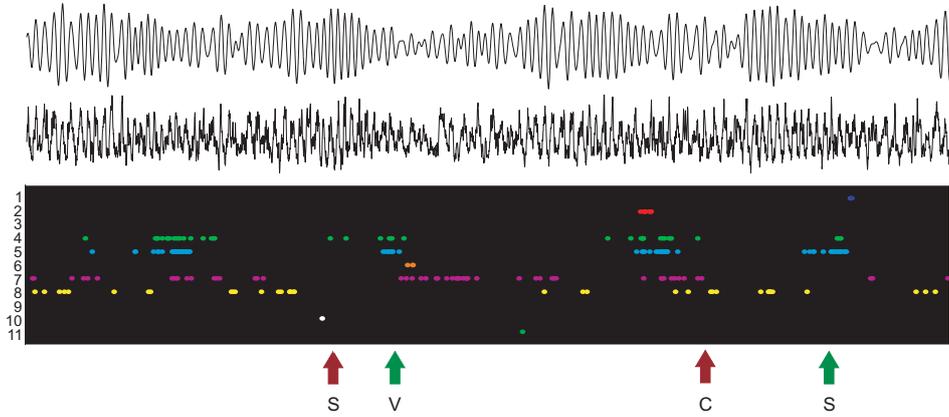
*Histology* The final position of the tetrodes was marked by passing a 25  $\mu$ A current lasting 10s through one lead of each channel to produce a small lesion. The next day, animals were transcardially perfused with a 0.9% NaCl solution followed by 4% paraformaldehyde in phosphate-buffered saline (0.1 M, pH 7.4) before the brains were removed. Coronal brain sections (40  $\mu$ m) were cut on a Vibratome and Nissl-stained for verification of tetrode tracks and end points. All of the tetrodes endings were in the ventral striatum approximately between 2.2 and 1.2 anterior to Bregma and between 1.6 and 3.0 laterally as compared to an atlas of the rat brain (Paxinos and Watson, 1986). To estimate the number of recordings originating from the core and shell subdivisions of the ventral striatum, we first assessed the endpoints of the individual tetrodes in the histological sections and converted them to coordinates according to the atlas (Fig 4.1B). We then calculated the approximate depths of the tetrodes in each session by subtracting the estimated travel distance from the tetrode endpoint. Thirty sessions yielded recordings from 203 locations of which 109 (54%) were likely in the core and 94 (46%) in the shell region. Note that ensembles from most sessions were likely to contain both core and shell recordings. Six recording sessions were identified as containing core-only recordings, and reactivation was present in these sessions.

91

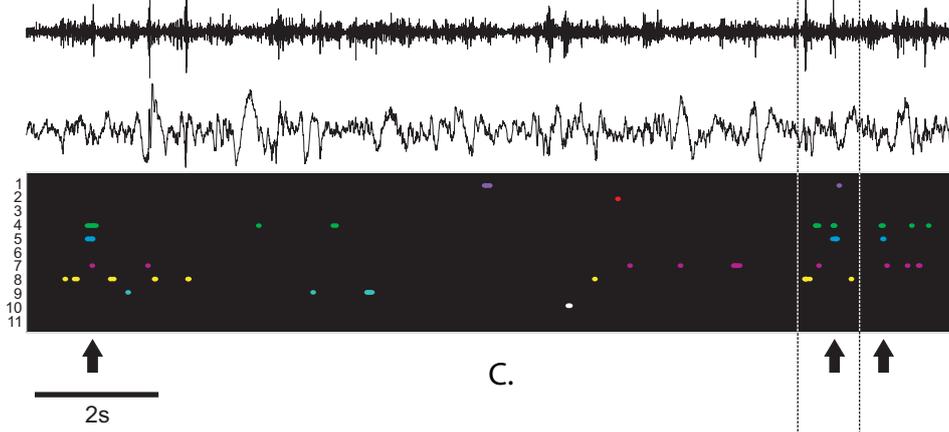
## Results

Activity of multiple single units in the VS was monitored in four rats during daily episodes of reward searching behavior on a triangle track flanked by two rest periods. On the track, rats typically ran from well to well and stopped each time checking for reward availability. As the rats became proficient at acquiring rewards on the track, the number of laps ran in twenty minutes increased over sessions ( $15.3 \pm 4.1$  in the first to  $62.3 \pm 6.2$  in the tenth, linear regression  $R^2 = 0.53$ ,  $p < 0.0001$ , see Supporting Figure 4.1A). The travel time between two wells was significantly longer when a reward was

A. Track running

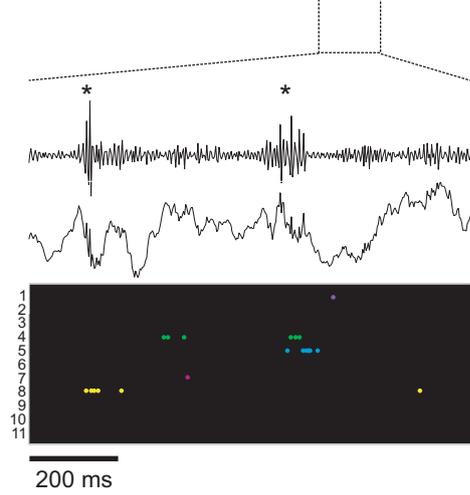


B. QW-SWS



92

C.



#### ◀ Figure 4.3

*Examples of the firing patterns of a ventral striatal ensemble, plotted in parallel with the local field potential recorded near the hippocampal fissure (A) and the pyramidal cell layer (B,C). A: As the rat is running on the track, theta oscillations were observed in the hippocampal field potentials. The bottom and top traces represented raw and filtered (6-10 Hz) field potentials, respectively. In a period of 15s, the rat crossed 4 reward sites (s = sucrose, v = vanilla desert, c = chocolate mousse), 2 of which were baited (indicated with green arrows) and two were not (red arrows). Each row in the plot below the field potential traces represents a single unit, its spikes being marked by specifically colored dots. Note the high variability in the firing patterns around each reward site arrival but also the positively correlated firing of the 'green' (4) and 'blue' (5) units. B: The hippocampal local field potential during QW-SWS is dominated by large irregular activity interleaved with sharp wave-ripple complexes. Bottom and top traces represent raw and filtered (100 – 250 Hz) field potentials. Note the concurrent firing of the 'green' (4) and 'blue' (5) units (two high contributors) indicated by black arrows. C: Enlargement of a segment of the LFP traces and spike patterns shown in B. The firing of units 4 and 5 as well as spikes of unit 8 are aligned to identified ripples in the hippocampal LFP which are marked with asterisks.*

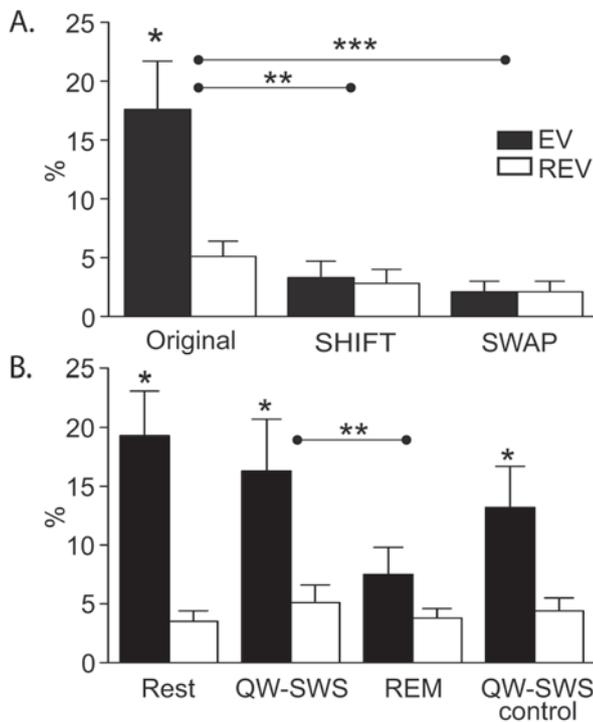
consumed than when an empty well was visited ( $16.54 \pm 0.38$  and  $7.65 \pm 0.25$  s, MWU,  $p < 0.0001$ ). The inter-well intervals did not become significantly shorter when several empty wells were encountered consecutively.

93

#### ***Behavioral correlates of ventral striatal firing patterns***

A total of 398 well-isolated, stable and sufficiently active single units was recorded over 30 sessions ( $13.3 \pm 0.9$  per session). Of these neurons, 79 (19.8%) showed significant firing rate changes correlated to reward site visits, whereas 7 (1.8%) showed changes to other task components such as locomotion. Responses time-locked to reward site visits were generated both by putative fast spiking interneurons ( $n=13$ ), which were not included in the analyses below, and other neurons (probably mostly medium-sized spiny neurons;  $n=66$  recorded from 25 sessions. Five sessions recorded from 3 rats did not contain any reward-related correlates). Neurons showing reward-related correlates exhibited various firing profiles (Fig. 4.2, Table 4.1). They responded predominantly by increasing their firing rates (61, 92.4 %), whereas occasionally firing rate decrements were observed (5, 7.6%). Individual responses peaked prior to the rat's arrival at a well (20, 30.3%), during the post-arrival phase (20, 30.3%) or consisted of elevations during both phases (26, 39.3%, see Supporting Figure 4.2). More neurons changed firing rates when a reward was present (39, 59.1%) than when it was absent (9, 13.6%). Five neurons (7.6%) showed firing rate changes in both conditions but with significantly different response magnitudes, whereas the remaining units showed no difference between presence and absence of

reward. About half of the neurons discriminated between reward sites, by either selectively responding to a single site (21, 31.8%), two sites (9, 13.6%) or to all three sites (6, 9.1%). In the latter two cases, responses to individual reward sites could differ in magnitude (9 out of 15 units). Interestingly, a subset of neurons (26, 39.3%) discriminated between the reward versus no-reward condition as well as between different reward sites. A last subset of cells (8, 12.1%) was less selective in their firing profiles in that their firing rate changes reached significance only when all reward conditions (absence/presence and sites) were lumped together.



**Figure 4.4**

*Reactivation in the ventral striatum: general characteristics and sleep phases. A: The EV exceeded the REV when compared across sessions (Original:  $***p < 0.001$ ,  $n = 30$ ). Reactivation disappeared when the temporal alignment between spike trains was disrupted (SHIFT,  $EV_{shift}$  versus  $EV_{original}$   $***p < 0.001$ ) or when cell identities were randomized (SWAP  $EV_{swap}$  versus  $EV_{original}$   $***p < 0.001$ ). B: Reactivation observed during QW-SWS sleep ( $**p < 0.02$ ) was not different from the reactivation found for the entire rest episodes*

*which included all episodes of motionless behavior (Rest,  $n = 16$  sessions). In contrast, REM sleep periods did not show reactivation and EV values were significantly lower than for QW-SWS ( $*p < 0.05$ ). This lack of reactivation cannot be ascribed to the relatively late occurrence of REM sleep periods after sleep onset because post-REM QW-SWS periods showed significant reactivation (QW-SWS control,  $*p < 0.05$ ). The strength of the observed reactivation during these intervals was not significantly different from the complete QW-SWS periods. Error bars indicate SEM.*

In line with previous studies on rat VS (Roitman et al., 2005; Tran et al., 2005), we will apply the term 'reward-related' to all units showing significant responses time-locked to reward site visits. Three units responded to one reward site only and in addition, did not fire differentially for reward presence versus absence. Although these cells were included in our analysis as reward-related, it cannot be excluded that their firing is purely spatially modulated (Lavoie and Mizumori, 1994; Shibata et al., 2001). However, exclusion of these units from analysis yielded similar results. When the firing patterns of reward-related and other units were viewed together during track running, sequences of consecutively firing neurons were observed, which however showed great variability because of time-varying allocation of reward to the three sites (Fig. 4.3A). Figure 4.3B shows the same ensemble firing in parallel with hippocampal local field potential during subsequent QW-SWS.

#### ***Reactivation in the ventral striatum***

The ventral striatum showed reactivation (EV:  $17.6 \pm 4.1\%$  and REV:  $5.2 \pm 1.3\%$ ,  $p < 0.001$ ,  $n=30$ , Fig. 4.4) which is in line with previous results (Pennartz et al., 2004). The EV exceeded the REV in a majority of sessions (23/30). Including putative fast-spiking interneurons in the analysis yielded similar reactivation values (EV:  $15.9 \pm 3.2\%$ , REV  $4.4 \pm 1.1\%$ ,  $p < 0.01$ ). When the correlations between cells recorded on the same tetrode were removed from the analysis, we still observed a significant reactivation (EV  $11.0 \pm 3.6\%$ , REV:  $3.7 \pm 1.0\%$ ,  $p < 0.05$ ). To examine whether the observed reactivation was attributable to cell- and time-specific firing correlations, EV and REV values were recalculated after the spike trains of the behavioral episode had been randomized in two different ways (Louie and Wilson, 2001). Both 1) randomly shifting these spike trains in time and 2) reassigning these spike trains to different cells strongly decreased EV values and removed the significant difference between EV and REV (1: EV:  $3.3 \pm 0.7\%$ , REV:  $2.8 \pm 0.2\%$ , 2: EV:  $2.1 \pm 0.9\%$ , REV:  $2.1 \pm 0.9\%$ , n.s. Fig. 4.4A). Reactivation measures in these control procedures were significantly lower than the original values ( $p < 0.002$ ). Furthermore, the strength of reactivation was significantly correlated with the progression through the sessions (linear regression on the difference (EV-REV),  $R^2 = 0.24$ ,  $p < 0.01$ ) and with the number of laps ran on the track ( $R^2 = 0.17$ ,  $p < 0.05$ ; See Supporting Fig. 4.1B). As the detectability of reactivation may positively correlate with the behavioral regularity and repetitiveness of task performance, this positive correlation per se does not confirm or contradict a role for reactivation in learning and memory consolidation (Jackson et al., 2006).

The dynamics of reactivation in the course of time after track running was studied by comparing two 20-minute blocks of concatenated quiet rest/sleep (block 1 EV:  $22.2 \pm 4.3\%$  REV:  $4.0 \pm 1.3\%$ ,  $p < 0.0001$ , block 2 EV:  $16.6 \pm 4.2\%$  REV:  $7.7 \pm 2.5\%$   $p < 0.002$ ). We

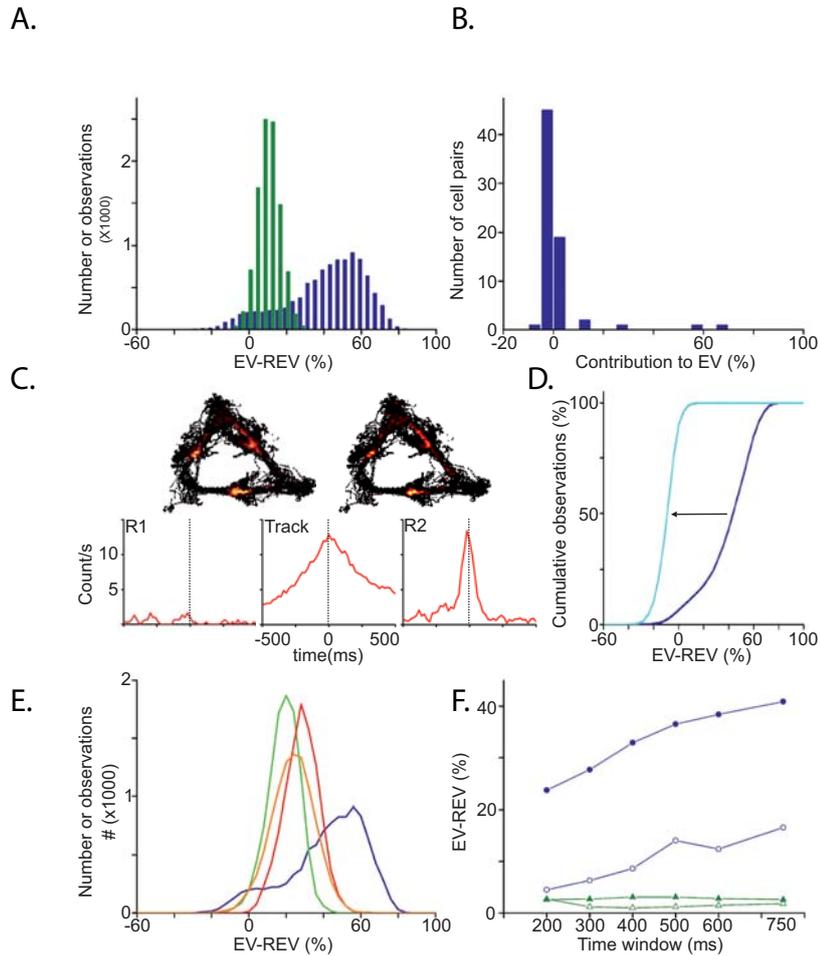
did not find a significant decline over the two blocks (n.s.). Because periods of awake behavior within post-experiential rest were excluded from this analysis, the amount of rest time in which rats slept for a total of 40 minutes was  $53.7 \pm 1.7$  minutes. Similar results were obtained when reactivation was computed over continuous 20 minute time blocks, taking only the quiet rest/ sleep periods within the blocks into account.

Ventral striatal reactivation occurs in SWS but whether it emerges during REM sleep has remained unclear (Pennartz et al., 2004). To address this issue, REM sleep and QW-SWS episodes were delineated in 16 sessions containing at least 4 minutes of REM sleep (pre-behavioral rest:  $8.0 \pm 0.7$ , post-behavioral rest:  $10.2 \pm 0.9$  min) and QW-SWS (pre-behavioral rest:  $20.4 \pm 2.0$ , post-behavioral rest:  $37.4 \pm 2.9$  min) per rest episode (Fig. 4.4B). Reactivation during QW-SWS was comparable in strength to the replay computed across whole rest periods in the same sessions (EV:  $16.3 \pm 4.7\%$ ; REV:  $5.1 \pm 1.6\%$ ;  $p < 0.02$ ). Specifically, reactivation appeared particularly strong during short time-windows (200 ms) following ripple-onset (EV:  $22.3 \pm 5.3\%$  and REV:  $2.8 \pm 3.8\%$ ;  $p < 0.001$ ) whereas for windows of identical length taken from the intervals between ripples the EV was not significantly different from the REV (EV:  $13.0 \pm 4.7\%$ , REV:  $8.8 \pm 3.1\%$ ). Reactivation for spikes within the ripple-windows was significantly different from spikes in interval-windows ( $p < 0.01$  for both EV and the difference EV-REV). We noted that the duration of ripples in post-behavioral QW-SWS was slightly but significantly increased compared to pre-behavioral QW-SWS ( $99.0 \pm 0.4$  ms and  $95.3 \pm 0.3$  ms respectively,  $p < 0.02$ ). The emission rate of ripples was similar in both pre- and post-behavioral QW-SWS (pre-behavioral SWS:  $0.65 \pm 0.01$  Hz and post-behavioral SWS:  $0.66 \pm 0.01$  Hz, n.s.). Significant reactivation was not detected in REM sleep (EV:  $7.5 \pm 2.5\%$ , REV:  $3.8 \pm 0.8\%$ , n.s.). When bin sizes larger than 50 ms were used to capture one or more theta cycles, similar results were obtained. To control for possibly confounding effects caused by undersampling of REM sleep or the timing of REM sleep throughout the resting period, reactivation was computed over QW-SWS episodes that were equal in length to REM episodes and followed these in time. These episodes also showed significant reactivation (EV:  $13.2 \pm 3.5\%$ ; REV:  $4.4 \pm 1.1\%$ ,  $p < 0.05$ ), similar to the amounts found for total QW-SWS time. The mean firing rate across all cells was significantly higher in REM sleep than in QW-SWS ( $0.34 \pm 0.02$  Hz and  $0.23 \pm 0.01$  Hz respectively). Correspondingly, the lack of REM sleep reactivation could not be ascribed to an undersampling in terms of total spike numbers (REM pre-behavioral rest:  $2165 \pm 269$ ; post-behavioral rest:  $2645 \pm 396$ ; QW-SWS pre-behavioral rest:  $1319 \pm 192$ ; post-behavioral rest:  $1836 \pm 231$ ).

### ***Reactivation of motivationally relevant information***

We next addressed the question whether reward-related information was specifically reactivated in the VS. Reactivation was assessed for the subset of reward-related units (RRU) and the subset of units recorded in the same sessions without such correlates (NRU, non-related units). The number of RRUs per session was generally low ( $2.6 \pm 0.3$ ,  $n=25$  sessions). Therefore, all cell pair based Pearson's correlations per episode were pooled across sessions and animals for each subset. This procedure requires at minimum 2 cells per session, and therefore fewer RRUs ( $n=57$ ) could be used than recorded throughout all 25 sessions. The RRU group showed an extremely strong reactivation (EV: 50.8%, REV: 0.3%,  $n=57$  cells) whereas the NRU group ( $n=166$ ) yielded a much smaller EV value, i.e. 14.6% and REV 2.3%. To assess significance, we applied a bootstrapping procedure with resampling of pooled correlation values. The distributions showed a significantly higher EV than REV in both subsets ( $p<0.001$ ), demonstrating reactivation in both ensembles (Fig. 4.5A). Moreover, the mean reactivation of the RRU subset was significantly higher than the NRU group as assessed from the distributions of the difference EV minus REV (EV-REV,  $p<0.001$ ).

The (EV-REV) distribution of the RRU group, however, was broad and bimodal (Fig. 4.5A). This suggests that cell pairs are unevenly contributing to the observed reactivation. To estimate the relative contribution of each cell pair to the session reactivation, a pair was excluded from the population after which EV and REV values were re-computed. The difference between the session EV minus the EV after pair exclusion represents the estimated contribution of that pair to the session EV. Only 5 of 70 cell pairs (7.1%) contributed more than 10% to their respective EV but there were no pairs that negatively contributed to the EV by more than -6% (binomial test  $p<0.05$ , Fig. 4.5B,C). These 5 highly contributing cell pairs were all recorded from different sessions across 4 rats. All but one of these 10 cells showed increased firing rates prior to the rat's arrival at reward wells. Generally, the rate maps and peri-event time histograms of two cells forming a highly contributing pair were similar, and their cross-correlograms showed a prominent peak around zero, indicating a high degree of synchronous firing during track running and during the subsequent rest period (Fig. 4.5C). When all of the highly contributing pairs were excluded from the RRU group, the pooled EV value dropped to 0.63% (REV: 7.5%), indicating that reactivation in the RRU group strongly depended on a small fraction of the RRU population (Fig. 4.5D). The incidences of high-contributors ( $\geq 10\%$  EV) in the NRU pairs and the pairs consisting of one RRU and one NRU (mixed pairs) appeared lower than in the RRU group (NRU: 7/1379 (0.005%) and mixed: 3/545 (0.006%)) and were not significantly different from the incidences of pairs that made



98

a large negative contribution to the EV (i.e.  $\leq -10\%$ ) in both groups. Reactivation was still present in the NRU group when the high-contributors were excluded from the population (EV: 8.8%, REV 2.3%,  $p < 0.001$ ), however, this residual replay was weaker than in the total NRU population.

In principle, differences in the number of cell pairs, average firing rates or correlation strength between spike trains during the behavioral experience between the subgroups could account for the observed difference in reactivation between the RRU and the NRU groups (Fig. 4.5E). To examine these possibilities, the RRU group (mean firing rate:  $1.3 \pm 0.2$  Hz, mean correlation strength:  $0.033 \pm 0.007$ ) was tested against three subsets of the NRU group that were matched for number of cell pairs and 1) were taken from the same

#### ◀ Figure 4.5

*Strong reactivation of ventral striatal units showing reward-related firing patterns. A: Bootstrap distributions of EV-REV values of reward-related correlates (RRU, blue) and non reward-related correlates (NRU, green). The RRU distribution was broad and bimodal, indicating inhomogeneous contributions of cell pairs to reactivation. B: Distribution of the contributions of individual RRU cell pairs to their session EV. Only a small fraction of the RRU cell pairs contributed more than 10% (5/70, 'high contributors'). C: Rate maps of two neurons constituting a highly contributing pair. Cross-correlograms show that firing patterns of these cells, recorded on different tetrodes, were highly correlated in time on the track and in post-behavioral rest, but not in pre-behavioral rest. D: Reactivation disappeared after exclusion of highly contributing pairs from the total RRU group as shown by a leftward shift in the cumulative EV-REV distribution (blue → light blue). E: Mean reactivation in the RRU group was a factor of 1.5 - 3.5 higher than in control groups matched for recording session (green), firing rate (orange) and strength of Pearson's correlation coefficients (red). F: EV-REV values plotted as function of the width of the time window before and after reward site visits or intervals. Spike patterns of RRUs (blue) that occurred in close temporal association to reward site visits (filled circles) were reactivated more strongly than those that occurred in intervals (open circles). Reactivation was not observed for the NRU group across the same time windows (green).*

sessions (session-matched NRU), 2) showed similar mean firing rate during track running ( $1.2 \pm 0.1$  Hz, rate-matched NRU) or 3) showed similar mean and distribution of Pearson's correlation coefficients ( $0.034 \pm 0.004$ , correlation-matched NRU). The NRUs comprising the various control groups showed similar waveform characteristics and inter-spike interval distributions as the RRUs. In line with the overall NRU group, reactivation was apparent in all three groups, most prominently in the correlation-matched NRU (EV: 33.6% en REV: 1.3%) and the rate-matched NRU (EV: 31.6%, REV: 5.5%) but also in the session-matched NRU (EV: 22.0%, REV: 0.4%, all groups;  $p < 0.001$ ). Reactivation in all three of these NRU subgroups however was not nearly as strong as in the RRU group ( $p < 0.001$ ).

The results presented thus far leave open the possibility that the reactivation observed for RRUs is not attributable to neural activity temporally linked to reward site visits but to intervals between them. Thus, we examined reactivation measures based on the narrow time windows during track running that were associated with reward site visits (-0.5 to +0.5s relative to site arrival), and found a significantly higher EV than REV when pairwise spike train correlations within these windows were compared to pre- and post behavioral rest episodes (EV: 36.9%, REV: 0.38%,  $p < 0.001$ , Fig. 4.5F). Reactivation occurred also in the intervals between these windows but was significantly less when either the total duration (EV: 14.6%, REV: 0.2%,  $p < 0.001$ ) or the spike counts (EV: 17.8%, REV 1.2%,  $p < 0.001$ ) were comparable to the reward-related window (reward site visits versus intervals,  $p < 0.001$ ). Time

windows ranging from 200 to 750 ms before and after reward site visits and corresponding intervals yielded similar results (Fig. 4.5F). The session-matched NRUs also reactivated both during the reward site visits (EV 19.4% REV: 6.6%,  $p < 0.001$ ) and intervals (EV: 10.6% REV: 0.9%,  $p < 0.001$ ). Although reactivation was stronger during the reward site visits ( $p < 0.001$ ), the difference was less prominent than for the RRUs. None of the other NRU groups showed the same pattern. In the overall NRU group, the rate-matched NRU group and the correlation-matched group reactivation was absent during both the reward site visits (overall NRU: EV: 1.2%, REV: 0.0%; rate-matched NRU: EV: 1.0%, REV 0.8%, correlation-matched NRU: EV: 2.3% REV: 0.8%) and the intervals (overall NRU: EV: 3.1%, REV: 0.0%; Fig. 4.5F, rate-matched NRU: EV: 0.2%, REV: 3.0%, correlation-matched NRU: EV: 0.1% REV: 0.1%). An increase of the time period flanking reward site arrivals and intervals of at least several seconds was required to obtain a reactivation strength comparable to that found for the whole track running period.

100

Next we tested whether the strong reactivation in the RRU group specifically occurred during a particular sleep phase. For QW-SWS, the EV (77.1%) greatly exceeded the REV (7.4%, MWU  $p < 0.001$ ). In contrast, for REM sleep low EV and REV values were found, 1.3% and 0.8% respectively (n.s.). The computation of RRU reactivation over QW-SWS episodes that were equal in length to REM episodes and followed these in time yielded an EV and REV of 55.4% and 0.1% (MWU  $p < 0.001$ , relative to REM sleep). Similar to the results for the complete population per session, this difference between QW-SWS and REM sleep for RRUs could not be explained by spike count (REM R1:  $759.6 \pm 138.9$ ; REM R2:  $991.2 \pm 204.4$ ; QW-SWS R1:  $407.1 \pm 107.3$ ; QW-SWS R2:  $710.1 \pm 155.5$ ), duration or temporal order effects.

In addition to the preferential reactivation of reward-related firing patterns, there may be a concomitant experience-dependent change in their temporal relationships with ripples. Therefore we considered a measure of the proportion of spikes occurring during ripple-windows or during intervals between ripple-windows, corrected for the different durations of these two types of state. The measure is also required to correct for differences in mean firing rate, because on average the RRUs generated higher rates than NRUs. This measure, termed Relative Spike Density, was computed as follows for ripple-windows and intervals between ripple-windows occurring during pre- or post behavioral rest episodes. First, the spike count was taken for the relevant state and episode under scrutiny (e.g. ripple-windows, pre-behavioral rest). This spike count was divided by the total spike count occurring in the entire episode under study (e.g. pre-behavioral rest). This ratio was next divided by the total period of time made up by the relevant state in the same episode (e.g. total duration of ripple-windows during pre-behavioral rest). Within the QW-

SWS periods of post-behavioral rest, the mean Relative Spike Density was significantly higher for RRUs during ripple-windows than during intervals ( $4.75 \times 10^{-4} \pm 0.20 \times 10^{-4}$  and  $2.24 \times 10^{-4} \pm 0.12 \times 10^{-4} \text{ s}^{-1}$  respectively,  $p < 0.0001$ ,  $n = 57$ ). In contrast, the Relative Spike Density during pre-behavioral rest was similar for ripple-windows and intervals ( $5.91 \times 10^{-4} \pm 0.55 \times 10^{-4}$  and  $4.76 \times 10^{-4} \pm 0.12 \times 10^{-4} \text{ s}^{-1}$  respectively, n.s.). The NRUs ( $n = 166$ ) on the other hand showed similar Relative Spike Densities for ripple-windows and intervals in both the pre- and post-behavioral rest episodes (pre-behavioral rest: Ripple-windows:  $4.93 \times 10^{-4} \pm 0.24 \times 10^{-4}$ , Intervals:  $4.91 \times 10^{-4} \pm 0.12 \times 10^{-4} \text{ s}^{-1}$ , ns, post behavioral rest: Ripple-windows:  $2.57 \times 10^{-4} \pm 0.14 \times 10^{-4}$ , Intervals:  $2.44 \times 10^{-4} \pm 0.07 \times 10^{-4} \text{ s}^{-1}$ , ns). Comparing the two cell groups during post-behavioral ripple-windows, the Relative Spike Density of RRUs was significantly higher than that of NRUs (MWU,  $p < 0.001$ ). During the intervals, however, an opposite tendency was expressed, viz. the Relative Spike Density of NRUs was slightly but significantly higher than of RRUs ( $p < 0.001$ ). These differences between RRUs and NRUs were not observed during pre-behavioral rest. Thus, not only are reward-related firing patterns more strongly reactivated relative to non-related patterns, reward-related patterns also become more temporally aligned to ripple episodes during post-behavioral rest, relative to pre-behavioral rest and to non-reward-related patterns.

## Discussion

101

Our main results indicate that reactivation mediated by reward-related neurons was significantly stronger compared to neurons not showing such activity and was particularly prominent for spike trains temporally linked to reward sites. A small minority of the units, pairs sharing highly similar firing patterns on the track, accounted for this reactivation, underscoring the sparseness of the phenomenon. Reactivation of reward-related neurons was prevalent during QW-SWS but could not be detected during REM sleep and did not decay across 40 minutes of sleep following the task.

Ensemble recordings from rat VS revealed firing patterns that were closely correlated in time to reward site visits. This result agrees with previous studies in rats and primates indicating that VS neurons fire in anticipation of reinforcement and in the post-reward phase (Schultz et al., 1992; Roitman et al., 2005). Indeed, no units in our study responded to a particular location on the track that was not directly related to reward. Pre-arrival responses were marked by upward ramps in firing rate, whereas post-arrival responses were usually sensitive to the presence of reward. In agreement with existing literature, the definition of reward-related firing behavior used here is a broad one including sensori-motor as well as valuation aspects of reward approach and consumption. Our first novel findings were that VS reactivation was associated with short time intervals following ripple onset, as previously

reported for hippocampus (Kudrimoti et al., 1999). Reactivation was not detected in REM sleep, and this contrast could not be ascribed to a potential undersampling problem for REM sleep or to a temporal arrangement during post-behavioral rest where REM sleep occurred later than QW-SWS, in combination with a decaying reactivation. As REM sleep appears important for consolidation of procedural memory (Gais et al., 2000; Stickgold and Walker, 2005), REM-replay may occur under different conditions than used here, e.g. in other brain structures than the VS (Maquet et al., 2000; Louie and Wilson, 2001) or during later sleep stages.

102

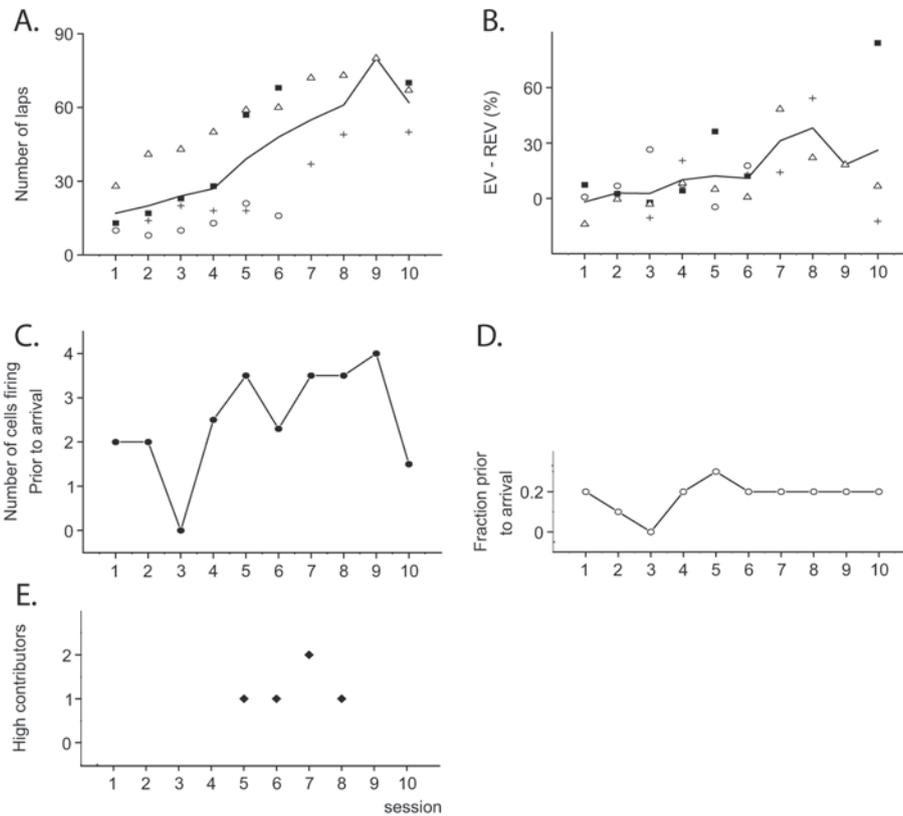
Our most striking observation was that the subgroup of RRUs reactivated very strongly and significantly more so than various control groups of NRUs. This difference could not be explained by differences in group size, mean firing rate or correlation strength. The powerful reactivation of RRUs was accounted for by only a small fraction of highly contributing cell pairs (7.1%), the members of which showed very similar firing behavior during track running (Fig 5C). This sparsity offers a parsimonious explanation of the previous (Pennartz et al., 2004) and current observation that reactivation in ventral striatal ensembles, although significant, is highly variable across sessions. Absence of reactivation in individual sessions is likely attributable to a lack of highly correlated cell pairs in the recorded sample. Since the reactivation remained strong when only spikes generated during a narrow time interval linked to reward site arrivals were included in the dataset from the behavioral period, our results indicate that at least reward-related firing patterns are reactivated, whereas other types of information present during the intervals between visits were reprocessed less prominently. What type of motivationally relevant information is reactivated is subject for further investigation, although the pre-reward firing of 'high contributors' suggests a component of reward expectancy.

The level of reactivation in the reward-related subpopulation was consistently higher than in the various non-related groups. Therefore it is tempting to speculate that the motivational value of the information processed during behavior might influence the strength by which it will be reactivated during sleep. Nonetheless, the subpopulation without significant reward-related activity also showed reactivation. The latter effect may have arisen because of reactivation of spike patterns correlated to behavioral components other than reward site visits, or because these cells showed a trend towards being correlated to rewarding events but remained subthreshold for statistical significance. Because NRUs may show a comparably strong reactivation as RRUs in another task setting than described here, further evidence would be needed to assess a hypothesized coupling between the motivational value represented by neural activity and its reactivation strength.

In contrast to the traditional view of the VS as a primarily “executive” interface that converts limbic, associatively learned information into goal-directed behavior (Mogenson et al., 1980), recent behavioral studies have highlighted its role in learning and memory (Kelley et al., 1997; Parkinson et al., 1999b; Corbit et al., 2001), acknowledging the long-term plasticity of intrastriatal glutamatergic afferents (Pennartz et al., 1993; Kombian and Malenka, 1994). Pharmacological studies using post-behavioral intrastriatal injections of protein synthesis inhibitors, NMDA or dopamine receptor antagonists suggest the VS to contribute to the consolidation of several forms of memory (Setlow and McGaugh, 1998; Hernandez et al., 2002; Sargolini et al., 2003; Dalley et al., 2005). Until recently, a possible neurophysiological mechanism explaining this involvement remained unknown. Given the current evidence that replay comprises a reward-related component, we hypothesize that the VS supports consolidation by endowing reactivation processes with motivational value, although it may also contribute cue- and context related information.

The empirical support for this thesis may have consequences for our current thinking about memory traces and their ‘off-line’ replay. Given the evidence that the hippocampus retrieves spatial-contextual information by reactivating cells characterized by place fields during active behavior, the VS appears to reactivate a different type of information. Hence it becomes reasonable to assume that both cortical and subcortical brain structures sustain replay in a domain-specific manner (Harris et al., 2001; Ji and Wilson, 2007). Whereas direct sensory processing of object information is generally accepted to occur in many brain areas in parallel and in a domain- or modality-specific way, so too may replay be conceived as a distributed process in which some structures reactivate information about the physical-sensory properties of an object and others about its spatiotemporal context and motivational value. It should be stressed that other brain structures involved in the processing of reward-related information including the medial prefrontal cortex and amygdala may influence or contribute to the reactivation of motivational information in the VS (Paré et al., 2002; Euston et al., 2007b).

Regardless of the precise nature of trace storage, reactivation across multiple brain structures such as hippocampus and VS raises the question how pieces of information belonging to the same scene or event are re-processed coherently, so that erroneous associations with other events can be prevented. Synchronized cross-structural replay (Qin et al., 1997; Hoffman and McNaughton, 2002; Ji and Wilson, 2007) may subserve such a function. As suggested by our data, hippocampal ripples may have a facilitating or coordinating role in supporting this process. Since ventral striatal replay especially occurs during hippocampal ripples and shortly thereafter, these high-frequency oscillations and the associated transitions in neocortical states (Battaglia et al., 2004b; Siapas et al., 2005)



may act as a release mechanism by which the principal cells of VS reactivate motivational information that will subsequently reach target sites in e.g. ventral pallidum and connected thalamo-prefrontal cortical circuits, lateral hypothalamus and ventral tegmental area (VTA). Recently, Foster and Wilson (2006) reported a reverse replay in hippocampal place cells when rats were pausing at a reward site after running along a track. They suggested that, when this replay coincides with decaying dopamine transients triggered by VTA activity, a reinforcement learning mechanism can be implemented to the effect that reward associations of places close to the reward site are more strongly stored than those of distant places. Such a process would require (i) reverse replay to be temporally aligned with transients in dopamine release, and (ii) the appropriate reward information to reach the VTA to enable it to produce a presumed teaching signal (Schultz et al., 1997). Hippocampal ripples triggering reactivation of ventral striatal principal neurons projecting to the VTA may provide such a temporal alignment mechanism (Pennartz et al., 2004; Foster and Wilson, 2006). The second requirement may be met by the current evidence for 'off-line' retrieval of value information, which may drive dopaminergic cells to signal errors in reward prediction. In turn, such signals may underlie dopamine dependent memory consolidation in target

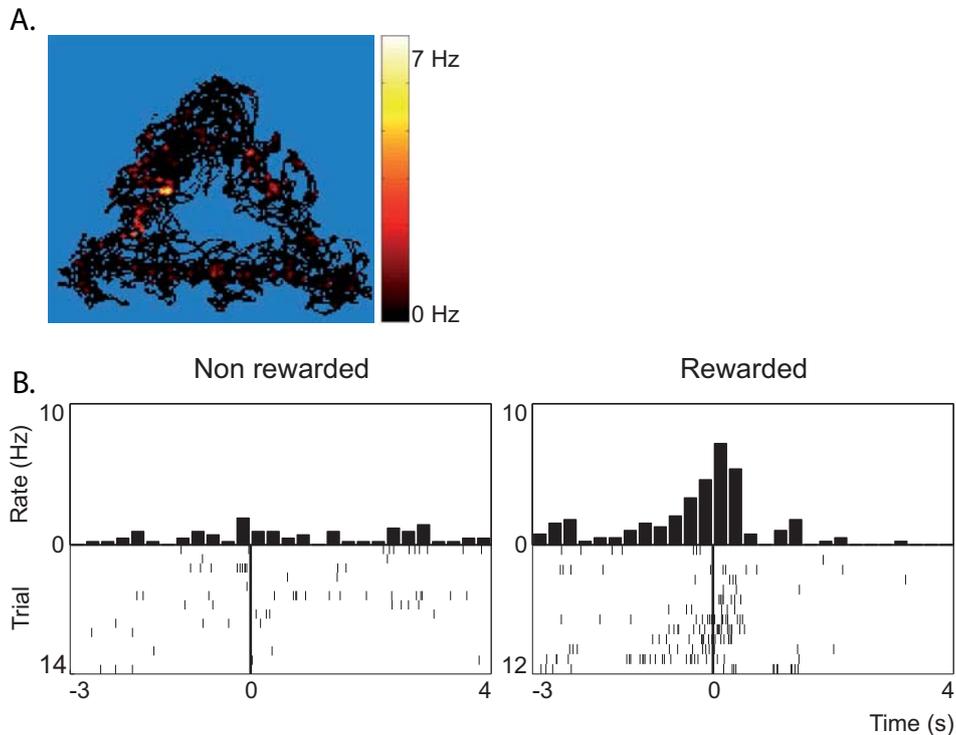
#### ◀ Supporting Figure 4.1

*A: The number of the laps completed by the rats was significantly correlated to the session number ( $R^2=0.53$ ,  $p<0.0001$ ) indicating that the rats became more proficient in acquiring rewards with experience on the track. The solid line represents the average number of laps ran by 4 rats. The symbols indicate the number of laps ran by the individual rats on each session. B: Reactivation expressed as the difference between (EV-REV) was highly variable but increased over sessions and was correlated to the session number ( $R^2= 0.24$ ,  $p<0.01$ ) and to the number of laps completed on the track ( $R^2= 0.17$ ,  $p<0.05$ ). The solid line represents the average difference (EV-REV) for 4 rats; symbols indicate individual observations. It might be that the increasing reactivation with higher numbers of laps reflects a learning process underlying improving task performance. However, it might also be that the detectability of reactivation is enhanced along with task familiarity because the behavior of the rat, and thereby the firing patterns of the neurons, become more regular with training, yielding more consistently repeated firing patterns. C: Plot of the average number of cells that fired prior to arrival at a reward site as a function of session number. The average ranged between 0 and 4 cells per session. D: The fraction of cells firing prior to reward site arrival (i.e., number of cells firing prior to reward site arrival / total number of recorded cells) did not increase with task familiarity. E: Reward-related high-contributor pairs were encountered from session 5 onwards. However, the sparse observation of high contributor pairs makes it impossible to draw firm conclusions on the relation between the incidence of high contributors and task familiarity.*

105

structures such as the VS (Dalley et al., 2005). The mechanism by which dopamine release in target structures can be boosted during 'off-line' processing is currently unknown. This could be mediated by burst firing of VTA neurons which however seems predominant during REM sleep but not QW-SWS (Dahan et al., 2007) or enhancement of local dopamine release by excitatory afferent input to target structures (Floresco et al., 2003).

Even if this hippocampal-ventral striatal-VTA loop (Lisman and Grace, 2005) subserves different functions than hypothesized, other VS output pathways, such as the mediodorsal thalamus-medial prefrontal cortex pathway (Zahm and Brog, 1992), are available as potential routes along which motivationally relevant information can be supplied to a network of structures involved in the consolidation of reward-dependent learning processes. VS reactivation can be hypothesized to contribute a motivational or emotional component to memory consolidation (McGaugh, 2000; Wagner et al., 2006) not only by strengthening intrastriatal input connections (Pennartz et al., 1993; Kombian and Malenka, 1994) but also by affecting reactivation in other (sub-) cortical networks.



106

### Supporting Figure 4.2

*Example of a neuron firing prior to and following reward site arrival recorded during a rat's first session on the track. A: Rate map of the neuron. The firing rate of the neuron is color coded as a function of the rat's head position on the track; the trajectory of the rat is shown in black. B: Peri-event time histograms for the rewarded and non rewarded conditions of the same unit. Different reward sites were pooled. The cell is responding in the condition when actual rewards were encountered. In the first few trials the local firing rate increases following arrival at the reward sites, but gradually the firing increments start earlier in time. This pattern suggests that pre-reward firing patterns are related to reward expectancy, developing quickly during initial task acquisition.*

### Acknowledgements

This work was supported by Human Frontier Science Program grant RGP0127/2001 to C.M.A.P. and B.L.M., Netherlands Organization for Scientific Research VICI grant 918.46.609 and Senter Novem BSIK grant 03053 to C.M.A.P. and MH046823 to B.L.M. We thank Jolanda Verheul for help with data analysis, A. D. Redish and P. Lipa for the use of cluster cutting software MClust and BBClust respectively, and our colleagues at the technology departments of the University of Amsterdam and the Netherlands Institute for Neurosciences. The comments on the manuscript by Francesco Battaglia, Jadin Jackson and Sander Daselaar are greatly acknowledged.



