Neural representation of reward information: coding by single cells and populations in rat orbitofrontal cortex

van Duuren, E.

Citation for published version (APA):

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.

Download date: 10 Mar 2020
Chapter 6

General Discussion
The main aim of the research described in this thesis was to examine whether and how reward-related parameters such as the magnitude and probability of reward are coded within the orbitofrontal cortex. To this end, ensemble recordings were performed while animals performed an olfactory discrimination learning task. Neural activity was examined to determine whether predictive information regarding the magnitude and probability of reward were coded by individual cells and populations of neurons, and to gain knowledge about how these parameters are actually represented within the orbitofrontal cortex in terms of specificity for trial phases, evolution over time within these specific trial phases, and to get insight in how these representations are built up as learning progresses. In addition, to be able to manipulate the activity of these orbitofrontal ensembles during behavior, the combidrive was developed.

This discussion will first provide an overview of the findings by chapter, followed by a discussion focusing on the contribution of the orbitofrontal cortex to reversal learning, since deficits in this type of learning are the most prominent feature after orbitofrontal damage. Finally, the results of the current thesis are integrated in a general view of orbitofrontal functioning during adaptive goal-directed behavior.

**Summary of results**

In Chapter 2, the coding of reward magnitude by single units in the orbitofrontal cortex was examined. The orbitofrontal cortex is known to be involved in the representation of the motivational significance of stimuli and in applying this information to the guidance of goal-directed behavior (Gallagher et al., 1999; Lipton et al., 1999; O’Doherty et al., 2003; Schoenbaum et al., 1999, 2003; Thorpe et al., 1983; Yonemori et al., 2000). Orbitofrontal neurons in primates were found to encode predictive information regarding upcoming reinforcers, demonstrating differential firing activity to various types or amounts of reinforcers (Hikosaka and Watanabe, 2000; Ichihara-Takeda and Funahashi, 2006; Padoa-Schioppa and Assad, 2006; Roesch and Olson, 2004; Roesch et al., 2006; Simmons and Richmond, 2008; Tremblay and Schultz, 1999; Wallis and Miller, 2003). In addition, findings in rat orbitofrontal cortex suggested predictive neural coding of both appetitive and aversive outcomes (Schoenbaum et al., 1998). However, it was not known whether reward-predicting information was quantitatively represented in the rat orbitofrontal cortex. This would be expected, since it is known that the magnitude of a primary reinforcer exerts a profound effect on the selection and speed of behavioral responses (Black, 1968; Bohn et al., 2003; Boysen et al., 2001; Brown and Bowman, 1995; Campbell and Seiden, 1974) and hence represents great motivational value. To examine whether information regarding reward prediction is quantitatively represented in the rat orbitofrontal cortex, ensemble activity was
olfactory discrimination ‘go/no-go’ task in which five different odor stimuli were predictive for various amounts of a rewarding sucrose solution or an aversive reinforcer, i.e. quinine. Ensemble recordings were made using an array of tetrodes (Gothard et al., 1996; Gray et al., 1995), a technique which provides the advantages of yielding high numbers of well-isolated cells. During the task, animals could obtain reinforcement after odor sampling by making a nose poke in the fluid delivery well. Upon the nose poke in the fluid well, there was a waiting period of 1.5 s, after which the reinforcer was delivered. During this waiting period, neural activity reflecting the expectancy for the upcoming reinforcement was predicted to occur.

The results obtained in this study showed that predictive information regarding reward magnitude is represented by single units in rat orbitofrontal cortex in multiple trial phases. During the task, animals were able to discriminate between the various amounts of expected reward (as visible by shorter response latencies for the larger amounts of reward), and neural correlates related to both actual and expected reward magnitude were observed. Responses related to the expectation of the upcoming amount of reward were found to occur within the waiting period prior to reinforcement delivery and during the execution of the behavioral response towards the fluid well after odor sampling (the ‘movement period’). About one-half of the neurons showing a behavioral correlate in these periods demonstrated differential firing towards the different reward sizes. They showed a variety of different tuning curves to reward magnitude, which is as compatible with coding in a parallel-distributed network as is a monotonic relationship between firing activity and reward magnitude.

To elaborate further on the neural coding of reward magnitude in the orbitofrontal cortex, Chapter 3 addressed the question whether this reward parameter could also found to be represented by the activity of orbitofrontal ensembles. Single cell studies can not show how predicted or actual rewards are dynamically represented at the population level in the orbitofrontal cortex, which is especially relevant for understanding how other, connected brain areas may read out population activity from this structure. Although two earlier studies described ensemble activity in the orbitofrontal cortex (Gutierrez et al., 2005; Schoenbaum and Eichenbaum, 1995b), it remained unknown how actual and predicted rewards are represented by ensemble activity within a specific trial phase during an operant task, and whether such a population code would be specific for different trial phases. Using two different reconstruction algorithms, the Bayesian method and template matching, the magnitude of reward could be decoded from population activity during the movement and waiting period and after reward delivery. The decoding score was only weakly dependent on the size of the neuronal group participating in the reconstruction, consistent with a redundant, distributed representation of reward information. Furthermore, decoding was found to be largely specific for trial periods,
meaning that to a fair extent the ensemble activity is specific for these particular trial phases, and carry-over of reward information to the next trial phase occurs only to a limited extent. Since animals learned to discriminate between at least some of the various amounts of expected reward during the task, the dynamics by which consistent coding develops as the learning task progressed was also assessed. The decoding performance was found to increase steeply across the first few trials of the behavioral session, an effect that could not be explained by a nonspecific drift in response strength across trials. Finally, when the population responses to quinine as negative reinforcement were compared to the appetitive sucrose reinforcement, ensembles were demonstrated to discriminate between these two different outcomes, meaning that coding in the delivery phase is related to the quality of reward. Altogether, these findings show that representations of reinforcer quality and magnitude are broadly distributed across ensembles with a high, sub-second time resolution.

In Chapter 4 we examined whether a different reward parameter, i.e. probability of reward, was represented in rat orbitofrontal cortex as well. As is the case for reward magnitude, the probability of reward is a key factor in decision-making as well. For example, in a process called probability discounting, the value of probabilistic rewards is downgraded as the reinforcer becomes more uncertain (Rachlin et al., 1991). Human brain imaging studies indicated a role for the orbitofrontal cortex in decision making under uncertainty. For example, Hsu et al. (2005) demonstrated a positive correlation between orbitofrontal activity and the level of unpredictability of reward. In addition, animals with orbitofrontal lesions preferred the larger, but uncertain reward, which is in accordance with the stronger risk-taking behavior as demonstrated by humans with orbitofrontal damage (Pais-Vieira et al., 2007). However, no neurophysiological evidence was available to demonstrate that single unit activity in the orbitofrontal cortex can actually code reward probability. To this end, ensemble recordings were performed during a probabilistic olfactory discrimination ‘go-no/go’ task, in which odors were now predictive of the probability of a pellet reward. The task design was similar to the task used to assess the coding of reward magnitude: after odor sampling, there was a waiting period of 1.5 s in the food trough, after which the pellet was delivered or not. The probability of expected reward was found to be coded in the orbitofrontal cortex in a fashion comparable to the magnitude of reward: during multiple trial phases, predictive information regarding the probability of reward was found to be represented by both single units and populations of neurons. During both the movement and waiting period, single unit activity was found to be modulated by reward probability, and this parameter could also be reconstructed from population activity significantly above chance level for both trial phases. Neurons that
discriminated between different expected probabilities showed a variety of different tuning curves. Of particular interest was the observation that a subset of neurons active in the reward delivery phase was found to respond specifically during unrewarded trials, and not during rewarded trials. These cells might signal the absence of reward when it is expected and hence may provide an error signal representing the violation of a positive reward prediction. Furthermore, we were able to demonstrate in this study that it was indeed probability of reward and not reward uncertainty coded by population activity - two distinct, but closely related reward parameters.

Finally, chapter 5 described a new technological development in the field of ensemble recordings. With the ability to record ensemble activity, additional questions arise, for example how neurotransmitters influence activity of cell populations. To gain more insight in the interaction between neurotransmitters or pharmacological agents and neural firing activity, the combidrive was developed, in which the technique of (reverse) microdialysis was implemented in a multi-tetrode array similar to those in use to perform ensemble recordings (Gothard et al., 1996; Gray et al., 1995). Hence, the combidrive combines a multi-tetrode array consisting of a circular row of 12 individually movable tetrodes and 2 reference electrodes to perform ensemble recordings with a movable and replaceable microdialysis probe to locally administer pharmacological agents.

Assessment of the combidrive showed that this device indeed allowed ensemble recordings simultaneously with reverse microdialysis, and that the firing activity of neurons in the prefrontal cortex was not affected by perfusion through the microdialysis probe per se. In addition, the combidrive was applied in a comparative study to examine the effects of cumulative concentrations of tetrodotoxin (TTX), lidocaine and muscimol on neural firing activity in the prefrontal cortex. Although these drugs are widely applied in behavioral studies to transiently inactivate selective brain areas (Albert & Mah, 1973; Brioni et al., 1989; Ivanova & Bures, 1990), until now little was known about the dynamics and reversibility of their inhibitory effect in relation to population activity in awake animals and the possible differences between them. The results of this comparative study showed that all drugs reduced neural firing in a concentration-dependent manner, but differences were observed in the extent to which firing activity of the population was diminished and in the speed and extent of recovery. Both muscimol and TTX caused an almost complete cessation of firing activity for longer periods of time at the highest concentration used, whereas the perfusion of lidocaine resulted in a smaller reduction of firing activity of the population and in the fastest recovery of firing activity after washout. The overall results from this study indicated that when during a behavioral experiment a longer lasting, but reversible inactivation is required, muscimol is the drug of choice. Although the effects of TTX were comparable to muscimol, the
latter has the additional advantage that it inactivates neurons locally, i.e. at the soma and the dendrites, whereas TTX also prevents the occurrence of action potentials in fibers of passage. Whenever an inactivation is required for a shorter period of time, lidocaine would be most suitable, although this drug was not able to fully block the population response. The results also implied that during behavioral studies it is important to take into account the variability in the neural response to the drug that is used, even in a confined region around the injection site. Furthermore, the combidrive turned out to be a useful tool to be applied in future studies to examine how neurotransmitters exert their effect on the activity of neuronal populations, as well as to clarify which neurotransmitters are actually involved in local cognitive or computational operations that take place in the brain during behavior, for example the contribution of serotonin or dopamine to processes of learning and attention.

The next question is whether the knowledge obtained as described in this thesis about the processing of reward-related information by orbitofrontal ensembles might provide a closer understanding of the involvement of the orbitofrontal cortex in reversal learning, in which subjects have to learn that previously rewarded stimuli are no longer rewarded whereas previously unreinforced events are now paired with reward. Impaired reversal learning is the most prominent, and probably most frequently reproduced feature of orbitofrontal damage (Bohn et al., 2003; Chudasama and Robbins, 2003a; Dias et al., 1996; Fellows and Farah, 2003; Ferry et al., 2000; Hornak et al., 2004; Izquierdo et al., 2004; Kim and Ragozinno, 2005; McAlonan and Brown, 2003; Meunier et al., 1997; Rolls et al., 1994; Schoenbaum et al., 2002, 2003b). Reversal learning itself was not experimentally examined in this thesis, since the first step in understanding the functionality of the orbitofrontal cortex should be to explore what kind of information is processed in this area and in what manner, which is most straightforward during task acquisition. The question that is central in the part below is how reversal learning depends on the orbitofrontal cortex, and we will explore how the results obtained in this thesis during task acquisition can add to the understanding of the functional role of this area in reversal learning.

The orbitofrontal cortex and reversal learning

Response inhibition

As already mentioned, damage to the orbitofrontal cortex results in the inability to rapidly learn reversals of previously acquired stimulus-reinforcer associations in reversal tasks. Initially, this impairment was considered to be caused by a failure in response-inhibition (Jones and Mishkin, 1972; McEnaney and Butter, 1969). In these studies, monkeys with orbitofrontal damage who performed an
object or place discrimination learning task demonstrated difficulties in reversing the previously acquired stimulus-reinforcer association. According to Jones and Mishkin (1972), the most plausible explanation for this deficit was a perseveration of the already established associations, although they did not exclude the possibility that this impairment was caused by difficulties in the formation of new stimulus-reinforcer associations. Similar observations came from more recent experiments that involved a variety of reversal tasks in multiple species (Bohn et al., 2003; Chudasama et al., 2003a; Dias et al., 1996; Fellows and Farah, 2003; Ferry et al., 2000; Izquierdo et al., 2004a; Kim and Ragozzino, 2005; McAlonan and Brown, 2003; Meunier et al., 1997; Rolls et al., 1994; Schoenbaum et al., 2002, 2003b). In most of these studies, a deficit in withholding responses became apparent after the reversal, but no difference between normal and lesioned animals was found in the acquisition of the initial discrimination, meaning that during this phase of the task animals learned to inhibit their responses at a similar rate as control animals. However, this observation is not accounted for in the explanation that the orbitofrontal cortex is involved in the inhibition of responding.

The idea that inhibition of prepotent responding is not a hallmark of orbitofrontal functioning is further supported by a number of more recent experiments. In a reversed reward contingency task, monkeys with orbitofrontal lesions were able to inhibit responses normally (Chudasama et al., 2007). In this task, animals selected a reward quantity by touching or reaching towards the rewards, without being able to get them. The selection of the smaller of two food quantities resulted in the receipt of the larger quantity, and vice versa. Reward magnitude represents strong motivational value, and the choice for the largest reward can be considered a strong prepotent response. However, orbitofrontal lesioned monkeys learned to suppress this natural response tendency in order to obtain the largest reward at the same rate as the control monkeys. Furthermore, Ostlund and Balleine (2007) showed that rats with orbitofrontal lesions were able to stop responding normally (i.e. stop pressing a lever) after reward devaluation during instrumental action-outcome learning in a lever-press task, indicating that the OFC is not critically involved in encoding the outcome of instrumental conditioning. Interestingly, testing of outcome-specific pavlovian-instrumental transfer (i.e. the facilitatory influence of pavlovian learning over instrumental performance) in the same task showed that orbitofrontal lesions made after, but not before training, abolished pavlovian-instrumental transfer (PIT). This finding provides additional evidence against the explanation for the deficits observed during reversal learning and reward devaluation paradigms (all tasks in which the target response depends at least partially on pavlovian conditioning; Roberts, 2006), that the orbitofrontal cortex has a general function in response inhibition: the effect on PIT was attributable to a failure to increase performance (i.e. increasing the number of lever presses in the
presence of a pavlovian stimulus), which is inconsistent with the idea of a deficit in response inhibition.

Altogether, these studies demonstrate that the orbitofrontal cortex is not required for the inhibition of responses, at least not in all situations. But when a deficit in response inhibition has to be ruled out as the cause of the observed impairments in reversal learning after orbitofrontal damage because this area is not critical for response inhibition, how does reversal learning then depend on the orbitofrontal cortex?

**Electrophysiology**

Schoenbaum and colleagues have provided electrophysiological data regarding orbitofrontal functioning during reversal learning in rats (see Schoenbaum et al., 2007 for review). The basic design of the dual-odor discrimination ‘go/no-go’ task used in these studies is that animals learn to make a ‘go’ response to the reward site to obtain reward after sampling an odor predictive of a reinforcing outcome (sucrose), but that this response will result in a punishment when made after sampling a second odor predicting an aversive outcome (quinine). Both control animals and animals with orbitofrontal lesions learn to respond to the positive odor, and not to respond after the negative one. During the reversal, in which the contingencies are switched such that the positive odor now predicts the negative outcome and the negative odor the positive outcome, lesioned animals are impaired in acquiring the new associations (Schoenbaum et al., 1999; 2003b). Performing electrophysiological recordings in both basolateral amygdala and orbitofrontal cortex showed that during this task, neurons within these two areas develop cue selective-activity (Schoenbaum et al., 1999; 2003b; see also Rolls et al., 1996; Thorpe et al., 1983), either in combination with a response towards the response outcome (referred to as cue-selective outcome-expectant neurons), or solely to the cue (referred to as cue-selective neurons). These neurons were furthermore found to either reverse their cue-selectivity during reversal learning, or to become non-selective. For example, during the initial learning phase, a particular neuron becomes selective to the positive odor, but after reversal the neuron switches its response to the previously negative odor, now that this odor predicts the positive outcome. Hence, these cells fire more in response to a stimulus predicting the positive response outcome. Neurons displaying the opposite pattern, which is responding more to the negative odor, are also found. Based on this finding, one could propose that the orbitofrontal cortex rapidly encodes reversals of cue-outcome associations. However, examination of the population response of the neurons displaying cue-selectivity in orbitofrontal cortex showed that the group of neurons referred to as cue-selective neurons actually loose their selectivity after reversal, and that after reversal a new cue-selective population develops (Stalnaker et al., 2006). Furthermore, it was
demonstrated in the same study that the probability of observing the reversal of cue-selectivity in the orbitofrontal cortex is actually inversely related to the rate of reversal learning, meaning that when more neurons are reversing their cue-selectivity, performance during reversal learning worsens. Hence, the flexibility of associative encoding in orbitofrontal cortex is inversely related to the speed of reversal learning. These findings indicate that flexible encoding in the orbitofrontal cortex is not an exclusive mechanism for reversal learning, since in that case one should expect that the population of neurons reverses cue-selectivity, together with an opposite relationship between performance during reversal learning and the probability of observing reversal of cue-selectivity, meaning a better behavioral performance with an increasing amount of reversing neurons.

An alternative hypothesis for the contribution of the orbitofrontal cortex during reversal learning was proposed by Schoenbaum et al. (2006). According to this hypothesis, the coding of expected outcomes within the orbitofrontal cortex may provide a signal that can be used to compare with the actual outcome and hence drive new learning, or adaptations of existing associations. When an actual outcome does not correspond with the expected outcome (which is obviously the case during reversal learning), this comparison can be used to alter, or update, the acquired representations. As a consequence, the orbitofrontal cortex would support reversal learning indirectly by facilitating changes in associative encoding in other brain areas, such as the basolateral amygdala, instead of directly, by rapidly encoding associations. However, one can object that during the acquisition phase prior to the reversal, in which stimulus-reward associations are presented that are unfamiliar to the animal, new learning occurs as well. In this respect, the proposal of Schoenbaum et al. (2006) still does not explain the ability of lesioned animals to normally inhibit responses during this phase.

Furthermore, it should be mentioned that the idea of the orbitofrontal cortex being responsible for reward driven learning was originally implemented in a computational model of reinforcement learning as proposed by Pennartz (1997). In this model, glutamatergic projection neurons of orbitofrontal cortex and basolateral amygdala are responsible for the processing of reward related information during reward-driven sensorimotor learning. Synaptic weights of these neurons come to reflect a stimulus-specific value of mean previous reward, and this reward value then acts as a predictor of future reward; when there is a constant performance upon the presentation of this specific stimulus, the mean reward value associated with this stimulus is also expected in future trials. The difference between the actual and the mean previous reward equals the error in the reward prediction, and these errors give rise to synaptic modifications in the sensorimotor network (i.e. the neocortex and striatum) that can drive new learning (Pennartz et al., 2000).
Relationship with the basolateral amygdala

As was demonstrated by Schoenbaum et al. (1999), neurons in the basolateral amygdala develop cue-selectivity as well, and, in contrast to the orbitofrontal cortex, this population does reverse its response after reversal learning (Stalnaker et al., 2007b). Hence, this population might be suggested to code the reversed cue-outcome associations. Furthermore, this associative coding was shown to depend on the orbitofrontal cortex, which is in line with the proposal that the orbitofrontal cortex would support reversal learning by facilitating associative coding in the basolateral amygdala. Cue-selective neurons in basolateral amygdala were unable to reverse selectivity in rats with orbitofrontal lesions (Saddoris et al., 2005) during reversal learning, which is probably caused by the absence of signaling of expected outcomes. Hence, the impairment observed during reversal learning seems to be mediated by the inflexibility of associative encoding in the basolateral amygdala, which is in turn caused by damage to the orbitofrontal cortex. Furthermore, when the orbitofrontal cortex supports reversal by detecting errors so that downstream areas can modify their representations, one should predict that lesions of the basolateral amygdala diminish the effects during reversal lesions after orbitofrontal lesions. In a recent study by Stalnaker et al. (2007a), it was demonstrated that lesions of the basolateral amygdala, which apparently do not have an effect on acquisition or reversal by themselves (Izquierdo and Murray, 2007, but see Schoenbaum et al., 2003a), correct the reversal impairment caused by bilateral orbitofrontal lesions. This suggests that the persistent coding of associations within basolateral amygdala, due to the absence of a correcting signal from the orbitofrontal cortex, impairs reversal learning. But still the issue remains how lesioned animals are able to withhold responses during the acquisition phase prior to the reversal.

Alternatively, we propose a different model for the neural mechanism underlying reversal learning, which is slightly analogous to the model for extinction learning as proposed by Quirk and colleagues (see for review Quirk and Mueller, 2008). During extinction learning, conditioned responding to a conditioned stimulus decreases when a reward is omitted. Extinction, like all other forms of learning, is thought to occur in three phases: acquisition, consolidation and retrieval, which are processes that each depend on a specific structure, namely the amygdala, the prefrontal cortex and the hippocampus. Based on experimental findings, it is hypothesized that the amygdala stores both memories for conditioning and extinction. Information about the conditioned stimulus enters the amygdala, hippocampus, and the infralimbic part of the medial prefrontal cortex (IL). The IL integrates the information about the conditioned stimulus with contextual information from the hippocampus in order to determine extinction retrieval. Furthermore, during extinction, the IL inhibits output from the amygdala to reduce fear. However, outside the extinction context, the output from the amygdala remains
uninhibited. The IL may therefore emit a ‘safety signal’ that can overrule a primary, prepotent response tendency elicited by a (formerly) fear-inducing stimulus (Milad and Quirk, 2002).

For the neural mechanism underlying reversal learning, we propose a model in which the primary association between stimulus and reward during reversal learning is formed within the amygdala, whereas after the reversal the orbitofrontal cortex is able to overrule this primary association. For example, in the acquisition phase of a conditioning task, primary associations between stimulus A predicting reward, and stimulus B predicting punishment, are formed in the amygdala. This associative information is projected to the orbitofrontal cortex, which also has the capability by itself to form stimulus-reward associations. After reversal, when the initial contingencies are altered, the primary association is overruled by the orbitofrontal cortex, since the alterations in the known stimulus-reward associations require an adaptation in behavioral responding. In this process, the primary association in the amygdala is not erased, but is still retained in memory (comparable to what happens during extinction learning; see Quirk and Meuller, 2008), and the orbitofrontal cortex guides the required alterations in behavior based on the secondary, higher-order association, a result of ‘meta-rule’ learning. This model also explains why animals with orbitofrontal lesions are still able to learn the initial associations during reversal learning, since the primary association is formed in the amygdala. Hence, the orbitofrontal cortex subserves a dual function in reversal learning: together with the amygdala it is involved in the formation of the stimulus-reward associations, but simultaneously it also controls activity within the amygdala. It is important to stress here that the primary associations can also be formed elsewhere in the brain, i.e. outside the orbitofrontal-amygdalar circuitry (e.g. in the medial PFC and ventral and dorsal striatum). Hence, orbitofrontal cortex and amygdala would not be the only brain regions involved in reversal learning, which does not contradict the hypothesis of the orbitofrontal cortex mediating higher-order learning per se. Confirming the non-exclusive involvement of orbitofrontal cortex and amygdala in reversal learning, it was previously demonstrated that monkeys with lesions of the rhinal cortex are impaired in object reversals, which is a type of reversal learning in which the identity of an object is predictive for future reward (Murray et al., 1998). Considering the fact that this is a different type of reversal learning that makes use of a visual discrimination task, whereas Schoenbaum and colleagues use an olfactory discrimination task, one might expect that this particular task may rely on different brain systems. Nevertheless, the rhinal cortex is known to have widespread connections with the neocortex and the amygdala (Insauti et al., 1997; Shi and Cassell, 1999), so there may well be a general role for this brain area in reversal learning as well.
Interesting in this respect is the recent finding that complete, bilateral lesions of the amygdala in the monkey did not impair object reversal learning in monkeys, but did impair performance in a reward devaluation task, which means that monkeys were unable to shift their choices for objects on the basis of changes in the associated reward value (Izquierdo and Murray, 2007). As suggested by the authors, these results indicate that the amygdala makes a specific contribution to the process of associating stimulus and reward: the amygdala would not be critical for guiding choices based on the representation of reward contingency (at least not in object reversal tasks), but for guiding choices after changes in reward value, as evident from the observed impairments in the reinforcer devaluation task. However, one can object that this reversal task, since it is a visual discrimination task, can also be learned by visually-based performance rules in which the reward provides information independent of its reinforcing value (Gaffan et al., 1985). Hence, learning in this task would then be independent of the formation of associations between stimulus and reward.

Single unit and ensemble activity: towards a unified electrophysiological view of orbitofrontal functioning

Results obtained with the electrophysiological single unit studies within orbitofrontal cortex and basolateral amygdala as described above did not involve population analysis of neural activity as described in the present thesis, but were based on the firing activity of individual cells. Although ensemble analyses can not clarify the role of the orbitofrontal cortex in reversal learning, they do provide additional knowledge about orbitofrontal functioning during olfactory discrimination 'go/no-go' tasks, in particular how dynamic the coding of reward-predictive information actually is. According to the proposal by Schoenbaum et al. (2006), and in agreement with the neural-network model by Pennartz (1997), the orbitofrontal cortex supports reversal learning by providing a signal that represents the discrepancy between the actual and expected outcome, which is subsequently used to alter associative coding in downstream areas. If this is indeed the case, rapid and flexible encoding of this information is required, since the ability to rapidly make decisions when environmental circumstances are changing can be vital. We demonstrated that the representation of reward predictive information by orbitofrontal ensembles is indeed flexible by nature, at least during initial task acquisition. Small subsets of orbitofrontal neurons coding reward predictive information are already sufficient for providing (parallel-distributed) signals that allow significant read-out in target structures, such as the basolateral amygdala. In addition, this information is available on short time scales (Chapter 4), which is another requirement to be able to rapidly make decisions. In the context of our alternative hypothesis for orbitofrontal functioning in reversal learning, i.e. exerting
higher-order control over the amygdala, striatum and other structures, a similar flexibility in forming secondary associations and in representing reward predictive information is required.

Of special interest is that reward predictive information was found to be coded in various stages in the process of decision making, including an early stage during which the ‘go’ response is executed (Chapter 2). The presence of predictive information in this phase supports a role for the orbitofrontal cortex in the control of conditioned responses and in encoding action-outcome associations. This is evident in reward devaluation tasks, in which animals with orbitofrontal lesions are unable to decrease conditioned responding to a conditioned stimulus according to an updated representation of the devaluated response outcome. This ability requires that predictive information regarding response outcomes becomes integrated with the decision towards and execution of the response, a process that may also require the orbitofrontal cortex (Baxter et al., 2000).

**Future perspectives and concluding remarks**

One aspect not included in this thesis concerns the activity of orbitofrontal ensembles during reversal learning. Since the orbitofrontal cortex and basolateral amygdala seem to form a complementary system involved in this type of learning, it would be of great interest to perform ensemble recordings simultaneously in orbitofrontal cortex and basolateral amygdala during reversal learning. Whenever an output signal from the orbitofrontal cortex is indeed necessary to update established associations in the basolateral amygdala (Schoenbaum et al., 2006), one would predict that the formation and alteration of these associations will occur with different time courses within these two structures. Furthermore, since several studies imply that more brain areas are involved in reversal learning, the next step in revealing the brain circuitry mediating reversal learning in its entirety could be recording neural activity in the rhinal cortex during this task, or inducing lesions in this area in combination with recordings in output areas, such as the amygdaloid complex. To test how generally our hypothesis of the orbitofrontal cortex mediating higher-order learning can be applied one can examine, when an animal has learned an association between a stimulus and a reward in a specific context, whether the orbitofrontal cortex is needed to learn that the same association is not valid in a different context.

In conclusion, the overall concept of population coding in the orbitofrontal cortex emerging from this thesis is that representations of various reward parameters, such as reinforcer quality, magnitude and probability are broadly distributed across ensembles and are characterized by a high, sub-second time resolution. The finding
that signals related to reward expectation are coded across different temporal phases along the process of decision making furthermore supports the hypothesis that orbitofrontal neurons collectively code a matrix of reward parameters as a function of the delay towards the moment of response outcome. Reward parameters are expressed more as modulatory signals rather than being main determinants of firing rate, and such a matrix provides the framework through which behavior can be altered upon environmental changes, regardless of whether the behavioral alterations are made voluntarily or not.

Although Schopenhauer concluded already in 1839 that man does not possess ‘free will’, he reached this conclusion solely by means of his own thinking, and without feeling the need of acquiring scientific evidence to support his thoughts and ideas. Nowadays, however, besides philosophers, neuroscientists are trying to solve this issue as well, seeking empirical evidence for the existence of ‘free will’. But the elucidation of the neural mechanisms of decision-making requires enormous effort, as hopefully became apparent from this thesis, and is only a small step towards solving the issue of how ‘free will’ is possibly exercised, especially when one considers that decision-making might be influenced by processes that do not even reach consciousness.

However, I do think that the work presented in this thesis will help in revealing the process of decision-making, and ultimately to solving the question of whether ‘free will’ can exist.