Emergence of Reward Coding in the Olfactory System

A dissertation presented

by

Daniel Joseph Millman

to

The Division of Medical Sciences

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

in the subject of

Neurobiology

Harvard University

Cambridge, Massachusetts

April 2016
Emergence of Reward Coding in the Olfactory System

Abstract

Identifying dangerous or rewarding elements in an animal’s surroundings is an important – if not primary – function of sensory systems. This holds particularly true for the mouse olfactory system since odors convey crucial information about predators, mates, kin and food. Thus, the olfactory system needs to effectively determine which odors are present as well as whether each odor has a positive or negative association, termed valence. Currently, we have little knowledge of how reward influences the processing of odors in the olfactory system of behaving mice. My work focuses on two high-level olfactory areas, the posterior piriform cortex (pPC) and olfactory tubercle (OT), that are situated at the intersection of sensory and reward-related brain regions. The pPC receives direct input from early olfactory areas and makes reciprocal connections to cognitive brain regions such as orbitofrontal cortex, limbic structures and the medial temporal lobe. The OT is a part of the ventral striatum which also receives input from early olfactory areas and is heavily interconnected with the reward system. To examine odor and reward coding in these areas, I developed a novel odor categorization task and recorded individual pPC and OT neurons during task performance. Mice successfully learn multiple, novel odor-response associations after only a few repetitions when the contingencies predict reward. I find that an explicit representation for reward category
emerges in the OT within minutes of learning a novel odor-response association, whereas the pPC lacks an explicit representation even after more than one month of overtraining. The explicit representation is visible in the first sniff of an odor on each trial, when the motor decision is made, and is not correlated with the trial-to-trial motor decision. Together, these results suggest that decoding of stimulus information required for reward-driven sensorimotor decision making does not occur within olfactory cortex, rather decoding occurs in circuits involving olfactory striatum.
Table of Contents

Abstract ........................................................................................................................................ iii

List of Figures .................................................................................................................................. vii

Acknowledgements ........................................................................................................................ x

1. Introduction and Background ........................................................................................................ 1
   1.1. Learning odor-reward associations ............................................................................................ 2
   1.2. Rodent sniffing behavior during experimental tasks ................................................................. 3
   1.3. Organization of the Mammalian Olfactory System .................................................................... 3
   1.4. The Piriform Cortex ................................................................................................................. 7
   1.5. The Olfactory Tubercle ............................................................................................................. 12
   1.6. The mammalian reward system and reinforcement learning .................................................. 16
   1.7. Reward coding in the olfactory system ....................................................................................... 17
   1.8. Odor-reward coding in high-level brain regions ..................................................................... 21
   1.9. Summary .................................................................................................................................... 22

2. Experimental Methodology ........................................................................................................... 23
   2.1. Experimental animals ................................................................................................................ 23
   2.2. Tetrode and nasal cannula implantation surgery ....................................................................... 23
   2.3. Odor stimulus delivery .............................................................................................................. 24
   2.4. Respiration recordings ............................................................................................................. 25
   2.5. Electrophysiological recordings and analysis ......................................................................... 25
   2.6. Behavioral training ................................................................................................................... 26
3. Odor and Reward Coding in pPC and OT ................................................................. 28

3.1. Behavioral task ................................................................................................. 30
3.2. Sniff behavior ............................................................................................... 38
3.3. Single unit responses of pPC and OT neurons .............................................. 39
3.4. Influence of sniffing on odor responses ......................................................... 46
3.5. Reward coding at the single-cell level in pPC and OT ................................ 51
3.6. Reward coding at the population level in pPC and OT ................................ 56
3.7. Multiplexing of odor and reward codes ......................................................... 71
3.8. Novel odor learning ....................................................................................... 74
3.9. Evolution of OT responses during learning of novel odor-reward associations .77
3.10. Analysis of error trials ................................................................................ 80
3.11. Summary of results ...................................................................................... 81

4. Conclusion and Discussion .................................................................................. 83

4.1. How do OT neurons acquire reward valence information? ......................... 83
4.2. How do OT neurons acquire reward-selective responses so early during learning? ............................................................................................................. 85
4.3. Is reward activity definitely absent in olfactory cortex (e.g. pPC) during decision-making? ............................................................................................................. 86
4.4. What gives rise to the disparity in lifetime sparseness between pPC and OT? 86
4.5. How to dissociate the contributions of reward valence versus motor decision in neural activity? ............................................................................................................. 87
4.6. Final thoughts .................................................................................................. 87
List of Figures

Behavioral task design ........................................................................................................... 31
Task performance .................................................................................................................. 33
Introduction of novel odors to the stimulus panel ............................................................... 35
Sniff behavior during task performance ............................................................................ 38
Histology example ................................................................................................................. 40
Single-unit responses of pPC and OT neurons ................................................................. 42
Population odor responses in pPC and OT ........................................................................ 44
Influence of sniffing on odor responses ............................................................................. 49
Reward valence selectivity during decision-making ......................................................... 53
Influence of odor sparseness on reward selectivity ......................................................... 55
Decoding reward valence from population data ............................................................... 59
Distributions of recorded cells by odor ............................................................................... 62
Training and testing linear classifiers on different sets of odors ...................................... 63
Reward and odor coding in OT ......................................................................................... 65
Reward and odor coding in pPC ......................................................................................... 66
Effects of sensory responsiveness on reward decoding ................................................... 69
Odor-by-odor relationship between behavioral performance and neural coding .......... 71
Multiplexing of odor and reward codes ............................................................................ 73
Learning of novel odor-reward associations ..................................................................... 76
Decoding reward selectivity from OT responses to novel odors ...................................... 78
Changes in OT cell responses to novel odors during learning ......................................... 79
OT selectivity on error trials ........................................................................................................... 81
Acknowledgements

First and foremost, I would like to thank Venki Murthy for excellent mentorship these past six years. He has built a lab filled with great scientists and all-around fantastic people. I feel incredibly privileged to have spent so much time in an environment with engaging discussion as well as the resources, expertise, and encouragement to explore questions that are interesting and important. Venki worked tirelessly to remove any barriers that might have stood in the way of us doing science that we found exciting. Thank you, Venki.

Of course, the Murthy lab would not have been so great without each and every person who has passed through over the years. I would especially like to thank Dan Rokni, who joined the lab at the same time as me. Working on the most closely-related project as we built up the behavior track in the lab, Dani was forced to listen to my endless ideas and complaints.

I learned a great deal from him.

I would like to thank the members of the Uchida and Dulac labs for making the hallway a fun place filled with concentrated expertise and diverse perspectives on olfaction.

Finally, I would like to thank you, the reader. I suspect fewer than ten people will ever feel motivated to read this far into the text. Congratulations and godspeed.
1. **Introduction and Background**

Identifying dangerous or rewarding elements in an animal’s surroundings is an important – if not primary – function of sensory systems. The sense of smell is the default means for many animals to identify and track (or avoid) distant objects and individuals as well as obtain life-or-death information about predators, mates, kin and food [Slotnick 2001]. Mice, like rats, have been a fruitful model organism to study olfaction, in part because these animals rapidly learn tasks based on odors [Slotnick 2001; Nigrosh 1975; Slotnick 1991; Rinberg 2006; Cohen 2012], which likely reflects their natural tendency to explore the world in the olfactory domain [Wesson 2008]. Rodents demonstrate a flexible ability to learn arbitrary odor-reward associations, which is a powerful tool for survival in their natural environments. Currently, we have little knowledge of how reward influences the processing of odors in the olfactory system of behaving mice. The work presented here focuses on two high-level olfactory areas, the posterior piriform cortex (pPC) and olfactory tubercle (OT), that are situated at the intersection of sensory and reward-related brain regions. The pPC receives direct input from early olfactory areas and makes reciprocal connections to cognitive brain regions such as orbitofrontal cortex, limbic structures and the medial temporal lobe [Datiche 1996; Haberly 1978; Johnson 2000]. The OT is a part of the ventral striatum which also receives input from early olfactory areas and is heavily interconnected with the reward system [Wesson 2011]. In this chapter, I give an overview of 1) rodent olfactory behavior in the context of reward learning, 2) the anatomy and physiology of the olfactory system with emphasis on the pPC and OT, and 3) the reward system and its role in reinforcement learning.
1.1. **Learning odor-reward associations**

Rodents are highly adept at learning arbitrary associations between odors, motor actions, and reward. A number of studies have demonstrated the ability of mice and rats to perform odor-cued operant behaviors, including odor discrimination [Schoenbaum 1995; Slotnick 2000; Uchida 2003; Abraham 2004; Rinberg 2006; Kepecs 2007; Barnes 2008; Doucette 2008; Gire 2013; Gadziola 2015], delayed match-to-sample [Lu 1993; Slotnick 2001], and paired-associate paradigms [Bunsey 1993]. Odor discrimination tasks have often employed a Go/No-Go paradigm in which the animals were required to decide whether the presented odor would be rewarded or not rewarded [Schoenbaum 1995; Slotnick 2000; Doucette 2008; Gire 2013; Gadziola 2015]. If the odor was associated with reward, the animal needed perform a motor action (Go) to receive the reward. Alternatively, if the odor was not associated with reward, the animal did not need to respond (No-Go). Mice and rats typically took a few sessions (<5) to learn the structure of the Go/No-Go task, then were able to reach near perfect performance within a single session with multiple novel odors [Slotnick 2000; Slotnick 2001]. The animals continued to successfully learn odor-reward associations with a new panel of novel odors each day for many consecutive days. Furthermore, the animals were not exposed to the odors prior to training and the reward valence of each odor was randomly assigned, indicating that the odor-reward associations were formed de novo. The reward contingency of some odors were then switched (from rewarded to unrewarded, or vice-versa) and the animals were able to successfully update their behavior accordingly within a single session. Therefore, rodents can rapidly learn multiple odor-reward associations, flexibly update previously-learned associations, and apply this information to make appropriate motor decisions. The next section will discuss how sensory information is acquired in olfaction in the first place.
1.2. **Rodent sniffing behavior during experimental tasks**

Rodents actively control the delivery of odors to sensory receptors through inhalation. When investigating a potential odor source, mice can modulate the amplitude and waveform of individual sniffs, and increase respiration from a basal rate of 3 to 5 Hz to rapid bouts of sniffing up to 12 Hz [Wesson 2008]. During odor discrimination tasks, rats and mice have been consistently found to require only a single sniff (<200ms) in order to gain enough sensory information to achieve maximal response accuracy [Uchida 2003; Abraham 2004; Rinberg 2006; Kepecs 2007; Zariwala 2012]. In anticipation of stimulus presentation, the animals would sometimes (but not always) sniff rapidly until the first odor-filled inhalation and then gradually return to basal respiration [Kepecs 2007; Wesson 2009; Doucette 2011]. Response accuracy and sensory neuron activation did not depend strongly on the amplitude or waveform of that first sniff [Kepecs 2007; Wesson 2008; Wesson 2009]. Given the task to discriminate an odor that predicts reward versus an odor that does not predict reward, the rate of sniffing did not differ between odors until more than one second of sampling [Doucette 2011]. The divergence in sniff rate occurred long after decision-making was complete but was closer to the time of reward delivery. Overall, the primary function of active sniffing during odor-guided tasks in the laboratory seems to be drawing a single sensory sample from the environment.

1.3. **Organization of the Mammalian Olfactory System**

At the first stage of odor processing, the main olfactory bulb (OB) receives monosynaptic input from olfactory sensory neurons (OSNs) in the nasal epithelium and sends a highly transformed output directly to cortex [Wilson and Mainen 2006; Mori and Sakano 2011; Murthy 2011]. Olfactory information is obtained from the environment through the binding of odorant molecules to an array of olfactory receptors (~1000 in the mouse). Many odorants activate a
given receptor and many receptors are activated by a given odorant; thus, odorant identity can be encoded by the combination of activated receptors [Hopfield 1999]. Early olfactory circuitry is largely segregated into distinct processing channels that are each dedicated to a single receptor. An OSN expresses only one olfactory receptor and innervates a glomerulus that receives input strictly from other OSNs that express the same receptor [Buck and Axel 1991; Ressler 1994]. In the mouse, each of the bilateral OBs contains approximately 1800 glomeruli, typically a pair of glomeruli per olfactory receptor [Mombaerts 1996]. The two types of projection neurons of the OB, mitral and tufted cells, receive excitatory input from only a single glomerulus. A variety of inhibitory neurons in the OB, which outnumber excitatory neurons by a factor greater than 100, shape the responses of mitral and tufted cells before sensory information is passed to the rest of the brain. A number of functional roles have been proposed for these interneurons, including normalization of responses to total sensory input [Olsen 2010], extending the dynamic range of responses across odor concentration [Olsen and Wilson 2008], decorrelating mitral cells within [Dhawale 2010] and across glomeruli [Arevian 2008], and competition across glomeruli to sparsify the odor representation [Koulakov 2011].

The output of the OB monosynaptically innervates no fewer than eight distinct regions in the rodent brain, namely the anterior olfactory nucleus (AON), tenia tecta, olfactory tubercle (OT), anterior piriform cortex (aPC), posterior piriform cortex (pPC), the nucleus of the lateral olfactory tract (nLOT), cortical amygdala, and lateral entorhinal cortex (LEC) [Neville and Haberly 2004]. The axons of mitral and tufted cells merge as they exit the OB to form the lateral olfactory tract, which runs from anterior to posterior along the ventral-lateral surface of the brain and innervates regions that are ipsilateral to the originating OB. The most anterior targets of the lateral olfactory tract (AON, aPC, and tenia tecta) are innervated by both mitral and tufted cells, with the exception of the OT, which appears to get input primarily, if not exclusively, from tufted cells. The axons of tufted cells do not extend beyond these anterior regions and therefore the
most posterior regions (pPC, LEC, nLOT, and cortical amygdala) are innervated exclusively by mitral cells. Whether a region receives input from mitral cells, tufted cells, or both might have important consequences because the content of the odor information differs between the two cell types. Tufted cells begin to respond to odors at lower concentrations and have a larger dynamic range than mitral cells, whereas mitral cells have substantially more numerous connections with OB interneurons and thus their output is more heavily modified than that of tufted cells [Mori and Sakano 2011].

In addition to parallel monosynaptic input from the OB to the various olfactory regions, serial connections between areas establish a largely feed-forward processing hierarchy in the olfactory system. At the lowest level of this hierarchy are the AON and aPC, which are strongly and reciprocally connected [Haberly and Price 1978; Hagiwara 2012]. Of these two areas, the aPC then provides the majority of feed-forward input to the pPC; feedback from the pPC to AON and aPC is negligible [Hagiwara 2012]. At the top of the processing hierarchy, the OT receives converging input from the aPC, pPC, LEC, and nLOT, but does not send projections back to any olfactory areas [Haberly and Price 1978].

Although the specific contributions of each olfactory brain region to behavior are not well understood, a number of distinct anatomical and physiological features suggest that some degree of specialization exists.

- The AON has an outer shell (known as pars externa) that serves as a relay for glomeruli in the two OBs with the same olfactory receptor to communicate [Yan 2008]. AON pars externa cells receive input from the ipsilateral OB and project exclusively to the granule cells near the corresponding glomerulus in the contralateral OB. This relay might function to compare olfactory receptor activation across OBs or reduce noise.
- The LEC is a part of the medial temporal lobe system involved in memory and spatial
navigation, but the LEC does not contain the grid cells for which the medial entorhinal cortex is famous. Instead, cells in the LEC seem to be involved in learning associations between objects and their location. For instance, some LEC cells respond only at locations where a particular object was present, even after the object has been removed [Tsao 2013].

- The layout of inputs from the OB to the cortical amygdala is stereotyped. The collection of axons from a single glomerulus innervates a small region of the cortical amygdala and that region is consistent across animals [Sosulski 2011]. This stereotyped pattern of OB input is not observed in other olfactory areas and might reflect hard-wired connections to drive innate responses to specific odors, such as those of predators.

- The piriform cortex, composed of the anatomically-distinct aPC and pPC regions, is a three-layer paleocortex and forms the largest part of the olfactory system. It is notable for having extensive recurrent excitatory connections, which have been suggested to allow flexible learning of novel odors as well as relationships among different odors [Haberly 2001; Wilson and Sullivan 2011].

- OT has anatomical features that resemble both cortex and striatum. Highly interconnected with the basal ganglia and the midbrain dopamine system, this area is believed to play a role in reward and reinforcement [Wesson and Wilson 2011; Giessel and Datta 2014].

The work presented here focuses on the piriform cortex and olfactory tubercle, so these areas are introduced in more detail in the following sections.
1.4. The Piriform Cortex

By far the largest target of OB input is the piriform cortex, which is a 3-layered paleocortex comprising 1) a dense sheet of excitatory neurons in Layer 2/3 whose odor tuning is determined by lateral olfactory tract (LOT) input and lateral connections among one another and 2) various interneurons that mediate feedforward and lateral inhibition that sharpens the tuning and duration of the odor responses of excitatory neurons [Zhang 2006; Young and Sun 2009; Luna and Schoppa 2008]. Piriform cortex is divided into anterior and posterior sub-regions, demarcated by a precipitous drop-off in the size of the LOT such that pPC receives relatively weak direct OB input compared to aPC. Recurrent connections within both aPC and pPC are non-topographic, thus non-local, and plastic, features that allow learning of novel odors and associations among odors. Connectivity between the aPC and pPC is not similarly promiscuous: aPC neurons project to the pPC but not vice versa, indicating that the relationship between aPC and pPC is that of a feedforward hierarchy [Hagiwara 2012]. pPC also integrates input from the rest of olfactory cortex, with the exception of the olfactory tubercle [Datiche 1996; Haberly 1978].

The entire piriform cortex projects to a variety of non-olfactory brain regions, including orbitofrontal cortex, amygdala, perirhinal cortex, and thalamus [Haberly 1978; Haberly 2001]. In fact, individual pPC neurons have elaborate axonal projections that branch off to many olfactory cortical areas (except aPC) as well as its non-olfactory targets [Datiche 1996; Johnson 2000]. Furthermore, the pPC is reciprocally connected with most of the non-olfactory regions that it targets, whereas aPC only receives strong feedback from orbitofrontal cortex [Lushkin 1983; Haberly 2001]. Together, the inputs and outputs of the piriform cortex suggest a rough model in which aPC more closely resembles a primary sensory cortex and pPC is more analogous to a secondary or association cortical area [Haberly 1978; Lushkin 1983; Johnson 2000; Haberly
Whereas the architecture of the OB is notable for the separation of glomerular processing channels, the local microcircuitry of the piriform cortex is well-suited to integrate arbitrary combinations of glomeruli. Although individual mitral cells only innervate a portion of the piriform cortex, simultaneously filling multiple sister mitral cells reveals that collectively the 10-20 mitral cells from the same glomerulus provide input to the aPC that spans its area [Ghosh 2011]. Conversely, retrograde labeling using a rabies-based strategy reveals that each piriform cortex principal neuron receives input from a subset of all available glomeruli without apparent spatial specificity within the OB [Miyamichi 2011]. LOT input synapses on the dendrites of piriform neurons in layer 1a, whereas association fibers from that form the recurrent connectivity among piriform cortical neurons synapse in layer 1b [Haberly and Bower 1989; Ketchum and Haberly 1993]. Three types of excitatory principle neurons of the piriform cortex are located in layers 2/3, and are arranged in a gradient such that superficial cells receive more feedforward input from the LOT whereas deeper cells receive more input from association fibers [Suzuki and Bekkers 2006; Suzuki and Bekkers 2011; Wiegand 2011]. Semilunar cells, so named for having apical dendrites but not basal dendrites, occupy layer 2a and receive strong input primarily from LOT. Superficial pyramidal cells in layer 2b have both apical and basal dendrites, and receive weak input from LOT and substantially stronger input from association fibers. The dendrites of deep pyramidal cells in layer 3 do not reach layer 1a LOT inputs and thus receive exclusively recurrent connections. Inputs from LOT are less plastic than associational connections within the piriform cortex due to A-type potassium channels that prevent back-propagation of action potentials into layer 1a [Johenning et al. 2009]. Restricting plasticity to associational connections might permit flexible learning of new odors without destabilizing responses to previously learned odors.

Among the neurons that receive LOT input are GABAergic cells in layer 1 that mediate
feed-forward inhibition. The input from these GABAergic cells sharpens the time course of principal neuron responses to olfactory sensory input by inhibiting firing approximately 20ms after the onset of excitatory input into the same principle neurons [Luna and Schoppa 2006]. Inhibitory neurons in the piriform cortex are more broadly tuned to odors than excitatory neurons, suggesting that this inhibition functions primarily to sharpen the timing of responses rather than modulate odor tuning [Poo and Isaacson 2009; Zhan and Luo 2010]. Finally, deeper inhibitory neurons in layers 2/3, which lack direct LOT input, provide lateral inhibition that is initiated by the inputs from piriform excitatory neurons, thus exhibiting feedback and lateral inhibition [Stokes and Isaacson 2010]. These inhibitory neurons project widely across the piriform cortex with no apparent spatial topography.

Perhaps given preference due to its size relative to other olfactory regions, nearly all studies of odor coding outside the OB have been in the piriform cortex, specifically the aPC. These studies reveal fundamental transformations in odor coding between the OB and piriform cortex. In the OB, each mitral or tufted (M/T) cell fires spikes in a particular temporal pattern through the course of a respiration cycle (i.e. a sniff), and odor tuning results from this temporal pattern being altered in a different way depending on which odor is inhaled [Dhawale 2010; Cury 2010; Macrides 1972; Sobel 1993; Friedrich 2004; Davison 2007; Chen 2009]. A given odor is thus encoded by the relative timing of spikes across the population of M/T cells, constituting a temporal “sniff-phase” code. In the piriform cortex, each cell receives input from M/T cells originating in many different glomeruli. This convergence provides the mechanism to compare the activation of different glomeruli, a necessary step to decode the identity of odors that activate multiple types of olfactory receptors. Indeed, the firing rate of cells in the piriform cortex are sensitive to the relative timing of M/T cell firing across glomeruli, thereby transforming a sniff-phase (temporal) code in the OB into a rate code in the aPC [Bathelier 2008; Miura 2012; Dhawale 2010; Cury 2010; Haddad 2013]. Cells that respond to a given odor are distributed
throughout the piriform cortex without any discernible topographical organization [Stettler 2009; Illig 2003]. Topography may be unnecessary as computational analyses indicate that odor identity can easily be read out from arbitrary combinations of aPC cells [Miura 2012]. The piriform cortex has been hypothesized to operate on principles of compressed sensing, in which a transformation from a low-dimensional sensory representation (e.g. 1000 olfactory receptors or glomeruli) to a high-dimensional representation (e.g. >100,000 piriform neurons) forms a robust representation of an originally high dimensional sensory space (e.g. the entire space of potential odorant molecules) [Ganguli and Sompolinsky 2012]. This holds true even in the case of random connections between low-dimensional and high-dimensional representations (e.g. LOT inputs from OB to piriform), alleviating requirements to generate specific wiring to deal with every stimulus that might become relevant to the animal. Whether or not downstream cells in the pPC then integrate aPC output to obtain tuning to categorical qualities of odors, such as a common valence or source, remains to be addressed.

Multiple neuromodulatory systems have been shown to modulate the activity of piriform cortex. Serotonergic input from the dorsal raphe nucleus inhibits spontaneous spiking of piriform neurons but leaves odor-evoked spike responses unchanged [Lottem 2016]; noradrenergic input from the locus coeruleus has been shown to have a similar effect [Hasselmo 1997]. This differential effect might serve to prioritize incoming sensory information over ongoing processing. Cholinergic input from the horizontal limb of the diagonal band (HDB) is thought to play a role in learning and memory in the piriform cortex. Learning of new associations between odors is impaired by lesions of the HDB or infusions of the cholinergic antagonist scopolamine into the piriform [Linster 2001]. The finding that acetylcholine enhances long-term potentiation at association fiber synapses in the piriform suggests a potential mechanism by which cholinergic input could facilitate learning [Hasselmo 1995]. Finally, dopaminergic input to the piriform is non-uniform, with the pPC receiving more heavy innervation than the aPC, but still far less than
neighboring olfactory regions, the olfactory tubercle and cortical amygdala [Datiche and Cattarelli 1996b]. The effects of dopamine on piriform activity or plasticity is not yet known.

Lesion studies demonstrate that the piriform cortex is necessary for a variety of complex odor-driven behaviors. The ability to learn to discriminate odors, or mixtures of odors, is unaffected by lesions that are restricted to the piriform cortex [Slotnick 1980; Slotnick 1985; Staubli 1987; Zhang 1998]. It is not clear whether the lack of an deficit is due to compensation by one or more of the many other brain regions that receive olfactory information from the OB, rather than non-involvement of the piriform in discrimination tasks. This issue is addressed in further detail in the following section titled “Reward Coding in the Olfactory System.” Piriform lesions do, however, impair olfactory-driven behaviors with higher cognitive demands than odor discrimination alone. For example, piriform lesions resulted in severe deficits in an odor delayed non-match-to-sample task that required working memory over the order of tens of seconds [Zhang 1998]; interestingly, ablation of the LEC did not cause a similar deficit despite providing olfactory input directly to the hippocampus. Similarly, piriform lesions led to an impairment in an odor-cued navigation task [Staubli 1987], which also requires working memory as well as the integration of odor and spatial information. Furthermore, selective lesions of the pPC disrupted the ability to recall remote but not recently learned odor-cued fear memories [Sacco 2010], suggesting that the pPC might be involved in the long-term storage of odor-valence associations. Together, these results suggest that the piriform cortex is involved in tasks that require the integration of olfactory and non-olfactory information or place demands on working or long-term memory.
1.5. **The Olfactory Tubercle**

The olfactory tubercle (OT) is located on the ventral surface of the brain, adjacent and medial to the piriform cortex [Wesson and Wilson 2011; Giessel and Datta 2014]. Together, the OT and the nucleus accumbens comprise the ventral striatum, which is part of the basal ganglia and has been implicated in reward, reinforcement learning and motor function. Anatomically, the cytoarchitecture of the OT resembles a hybrid of the cellular organization found in cortical areas and the basal ganglia. The ventral portion of the OT, nearest to the surface of the brain, is a 3-layer cortex of densely packed cells that contains primarily (>90%) D1- and D2-type dopamine receptor-expressing medium spiny neurons (MSNs) similar to those that constitute the majority of neurons in other regions of the basal ganglia. This dense cell body layer (DCL) folds in and out to form characteristic crests and valleys; the crests extend toward the pia and are visible on the ventral surface of the brain as cell-dense spots scattered throughout the approximately 5mm$^2$ of each hemisphere that is occupied by the OT [Wesson and Wilson 2011]. The area that is dorsal to the DCL (i.e. deeper from the ventral surface) is populated by a collection of 10-20 nuclei known as the Islands of Calleja (IC) [Giessel and Datta 2014]. Each island contains granule cells that project to the nucleus accumbens as well as the piriform cortex and amygdala [Giessel and Datta 2014]. Locally, the axons of MSNs in the DCL innervate the granule cells of the IC; the extent of lateral connectivity among MSNs is not known. Finally, intermingled with the IC in the dorsal OT are protrusions of a neighboring basal ganglia structure, the ventral pallidum. As a result of these protrusions, the border between the dorsal OT and the ventral pallidum is essentially non-existent at the anterior pole of the OT but becomes progressively more clear (and is ultimately distinct) toward the posterior pole of the OT.

Similar to the piriform cortex, sensory input from the olfactory bulb (OB) is carried via the fibers of the lateral olfactory tract. These inputs target the dendrites of MSNs in the DCL, with
preferential targeting of neurons in the crests of the DCL [Sosulski 2011]. Furthermore, there is a spatial bias for more dense innervation of the lateral OT, which is proximal to the lateral olfactory tract, and less innervation of the more distal, medial OT [Schwob and Price 1984; Wesson and Wilson 2011]. In contrast to the case of mitral cells providing the majority of OB input to the piriform cortex, the OT receives input primarily from OB tufted cells [Wesson and Wilson 2011; Scott 1980]. Tufted cells respond to odors at lower concentrations and generally have broader odor tuning than mitral cells, owing in part to fewer connections with inhibitory granule cells of the OB [Mori and Shepherd 1994]. In addition to direct input from the OB, the OT also receives indirect olfactory input from other cortical regions, including the piriform cortex [Johnson 2000; Carriero 2009]. Despite the apparent specialization for processing olfactory information, the OT might play a prominent role in multisensory integration as well. Approximately one-fifth of OT neurons recorded under anesthesia responded to auditory input [Wesson and Wilson 2010], although it is not clear whether auditory information is conveyed by monosynaptic connections from auditory cortex or arrives indirectly through a different part of the basal ganglia [Wesson and Wilson 2011].

Diverse and abundant neuromodulatory inputs suggests that sensory processing and learning in the OT occur in a state-dependent manner. Dopaminergic input from the ventral tegmental area (VTA) can have effects at the level of the DCL as well as IC [Giessel and Datta 2014]. Each MSN in the DCL expresses either the D1- or D2-type dopamine receptor, in roughly equal proportions, like their counterparts in the dorsal striatum [Le Moine and Bloch 1995; Gerfen and Surmeier 2011]. In slice experiments, stimulation of olfactory (LOT) inputs paired with phasic activation of VTA dopaminergic inputs resulted in potentiation strictly of the LOT fibers that were stimulated; further activation of the same LOT inputs in the absence of dopamine input extinguished the potentiation [Wieland 2015]. This result is a proof-of-principle that VTA dopamine input to the OT can act as a credit assignment signal that serves to identify
those olfactory stimuli that coincided with a salient event such as receiving reward [Reynolds and Wickens 2002]. Moreover, the granule cells in the IC are connected by gap junctions and this coupling strength is modulated by dopamine [Halliwell and Horne 1998]. Together with MSNs, this suggests that the vast majority, if not all, of the output neurons of the OT are affected by dopamine.

OT also receives substantial serotonergic inputs from both the dorsal and ventral raphe as well as noradrenergic input from the locus coeruleus [Wesson and Wilson 2011]. Raphe serotonergic neurons are reciprocally interconnected with VTA dopaminergic neurons, and have similar phasic responses to reward and punishment as the dopaminergic neurons along with slower changes in firing rate that seem to correlate with emotional states [Cohen 2015], though whether serotonergic input is involved in reinforcement is still an area of active research [Liu 2014; Fonseca 2015]. In addition to long-range neuromodulatory inputs, the OT has local cholinergic interneurons that are found throughout the dorsal and ventral striatum [Armstrong 1983; Millhouse 1984] Indeed, staining for acetylcholinesterase reveals striking contiguity with the rest of striatum and clearly distinguishes the OT from neighboring olfactory regions, such as the piriform cortex. Striatal cholinergic neurons seem to increase neurotransmitter release from neighboring dopaminergic terminals [Threlfell 2012; Nelson 2014].

The output targets of the OT parallel those of the dorsal striatum, but are not identical. For example, the primary target of MSNs in both areas is pallidum; however, MSNs in the OT project to the ventral pallidum whereas MSNs in the dorsal striatum project to the globus pallidus (i.e. dorsal pallidum) [Millhouse 1984]. The projection from the dorsal striatum to the globus pallidus is well known for the segregation of (1) a direct pathway in which D1-receptor-expressing MSNs monosynaptically output to the globus pallidus pars interna and (2) an indirect (disynaptic) pathway in which D2-receptor-expressing MSNs project to the globus pallidus pars externa which, in turn, projects to the pars interna. In contrast, projections from D1- and D2-
receptor-expressing MSNs of the OT to the ventral pallidum are not segregated [Smith and Kieval 2000]. GABAergic neurons in both parts of the pallidum then project to the thalamus (though targeting different nuclei), subthalamic nucleus, and the substantia nigra [Haber 1985; Zahm 1996]. Moreover, the dorsal striatum and OT provide a comparable amount of monosynaptic input to VTA dopaminergic neurons (by density, not total number of synapses; the dorsal striatum is substantially larger than the OT) [Watabe-Uchida 2012], thus also sharing the feature of strong reciprocal connectivity with the reward system. In fact, the OT is the largest source of input to the VTA from the olfactory system, providing a greater number of inputs than all other olfactory brain regions combined. Finally, OT projects to the lateral habenula [Fallon 1983], an area that conveys negative reward signals through inhibition of VTA dopaminergic neurons [Matsumoto 2007].

Experiments to determine the contributions of OT to behavior have not yet converged on a particular functional role; instead, these studies describe a variety of interesting phenomena to motivate future work. Despite ubiquitous dopaminergic input throughout the basal ganglia, rats have a higher drive to self-administer cocaine into the OT than either the ventral pallidum or the nucleus accumbens [Ikemoto 2003]. Likewise, conditioned place preference is also stronger with cocaine injections into the OT than the ventral pallidum or the nucleus accumbens. These results suggest that the OT might play a prominent role in reward and reinforcement. Broadly, OT lesions result in reduced behavioral engagement, such as decreased responsiveness to sensory stimuli [Hagamen 1977], locomotion, copulation and food consumption [Hitt 1973; Koob 1978]. In humans, fMRI measurements found that the OT was more active when odors were attended than unattended, an effect that was not observed in the piriform cortex [Zelano 2005]. Taken together, the anatomical and behavioral evidence point to the OT as positioned at the intersection of olfactory sensory processing and reward.
1.6. The mammalian reward system and reinforcement learning

The reward system is a group of interconnected brain regions, including the basal ganglia, thalamus, orbitofrontal cortex, amygdala, and midbrain dopaminergic neurons [Schultz 2000, for review]. Collectively, neurons in these areas encode multiple variables that are necessary for the learning and execution of reward-motivated behaviors. The most common signal conveyed by neurons in the reward system occurs when the animal has detected or received something of intrinsically rewarding value. Responses to reward delivery have been observed in dorsolateral prefrontal cortex, orbitofrontal cortex, amygdala, lateral hypothalamus, anterior cingulate cortex, substantia nigra, subthalamic nucleus, and all parts of the striatum. Many neurons, particularly those in the orbitofrontal cortex and amygdala, discriminate between reward and punishment as well as different magnitudes of reward [Thorpe 1983; Pratt 1998; Tremblay 1999; Schultz 2000]. Furthermore, stimuli that are not intrinsically rewarding but serve as cues that reward is likely forthcoming can also evoke responses throughout the reward system. Although these neurons discriminate between stimuli based on their relative associations with reward, unlike sensory cortex the responses are typically not sensitive to specific stimulus features [Tremblay 1999]. Reward-predicting cues establish an expectation for reward at a future time. When the expectations for reward are not met by reality, some neurons signal the discrepancy, known as reward-prediction error. The prototypical examples of which are midbrain dopaminergic neurons in the ventral tegmental area [Schultz 1997].

Reinforcement learning is a process through which animals can identify features of the environment and behaviors they execute that lead to future reward. Together, the combination of signals present in the reward system has been hypothesized to implement related models of reinforcement learning, such as Rescorla-Wagner, Delta Rule, and Temporal Difference (TD) learning [Sutton and Barto 1998; Schultz 1998; Lee 2012]. These models share the common
feature of attributing a reward value to stimuli, actions, or events that predict future reward, regardless of their own intrinsic reward. This strategy allows the animal to discover new paths to rewards and optimize their behavior to receive maximal reward. Predictors of reward can even be chained together to create sequences, such as complex motor behaviors, that lead to reward despite no individual members of the sequence ever being explicitly rewarded. The generality of this strategy is exemplified by the power and flexibility of classical and instrumental conditioning paradigms [Schultz 1998; Schultz 2000]. The reward prediction error is a key component of the learning algorithm, as the discrepancy between our predictions and the actual outcome tells us in what direction to adjust our predictions and by how much. As a consequence, midbrain dopaminergic neurons are believed to play an essential role in reinforcement learning. Not only does the activity of these neurons signal reward prediction error, but also dopamine can gate synaptic plasticity [Wickens 1996; Keiflin and Janak 2015], providing a biological mechanism to implement the update given by the error signal. The site of reinforcement that underlies learning of odor-reward associations, and whether learning depends on dopaminergic input, is not known.

1.7. Reward coding in the olfactory system

Neural activity that reflects the valence of odors has been reported throughout the olfactory system, even as early as the OB [Schoenbaum 1995; Doucette 2008; Doucette 2011; Gire 2013]. In these studies, rodents were trained to associate natural or synthetic odors that the animals were not likely to have previously experienced with a positive reward, such as water or sucrose, a negative reward (i.e. punishment), such as quinine or air puff, or no reward. In order to study odor-reward associations that were flexible and generic, the reward assigned to each odor was not related to the smell or molecular properties of the odorants. Following
training on odor-reward associations, the valence of odors had a significant but slow effect on cells in the OB and aPC. As expected for neurons in early sensory areas, the cells exhibited odor-selective responses that were time-locked to the 4-12 Hz sniffing that sampled the sensory stimulus [Doucette 2008; Gire 2013]. In addition, the firing rate of these cells was modulated on the timescale of seconds in a valence-specific manner. Thus, sensory and reward information is multiplexed in the responses of neurons even at the earliest stages of the olfactory system.

A few points should be kept in mind when considering the implications of these results. First, the effect of valence on the firing rate was much slower than the few hundreds of milliseconds that would typically be required for rodents to detect odors and make decisions based on the valence associated with those odors. At this long timescale, valence-specific activity could have been due to effects that were unrelated to the initial decision making process to determine the valence of an odor. The change in activity could have reflected post-decision behavioral changes, such as increased respiration as a side-effect of reward anticipation [Wesson 2008]. Second, the OB receives substantial feedback from multiple higher olfactory brain regions and neuromodulatory systems. The presence of reward-related activity in the OB could have resulted from feedback, learned changes to feedforward processing, or both. Finally, the number of odors used in these experiments was relatively small: eight or fewer in all cases. Since the responses of OB and aPC neurons are odor-selective, the average response of a single neuron to the couple of odors with one reward contingency might differ from the response to the couple of odors with a different reward contingency due to odor tuning alone.

To dissociate neuronal responses to a specific sensory stimulus versus the reward valence associated with that stimulus, a common paradigm is to train animals on (at least) two stimulus-reward contingencies, one positive and one negative, and subsequently switch the contingencies. If the neural response to each stimulus remains the same following the contingency switch, this is evidence that the neuron is selective for the stimulus and not the
associated valence; on the other hand, if the neural response switches with the contingency switch, this is evidence that the neuron is selective for the reward valence and not the identity of the stimulus. In such an experiment in the aPC, odor responses did not change when reward contingencies were switched [Roesch 2007], suggesting that aPC neurons had purely sensory responses (that is, under these specific behavioral conditions). In pPC, however, odor responses did change but in an unusual manner. The responses of pPC neurons did not purely follow the sensory stimulus or the reward [Calu 2007]. Instead, cells that had responded to a particular odor-reward contingency would neither respond to the same odor associated with the opposite reward valence nor respond to a different odor associated with an identical reward. Conversely, neurons that were previously unresponsive to any odor-reward contingencies would begin to respond to one or more of the same odors once those odors established a new reward association. This physiological evidence that pPC neurons are sensitive to the conjunction of sensory information and reward fits well with the anatomical evidence that pPC is at the intersection of sensory and cognitive neural circuits.

A recently published study that implemented this paradigm was the first to describe responses of OT neurons to olfactory stimuli in behaving rodents. The authors trained mice on a pair of odors, one of which cued an upcoming water reward while the other odor did not predict reward. They found that a majority of OT neurons were excited more strongly by rewarded odors than unrewarded odors [Gadziola 2015]. Upon reversal of the odor-reward contingency, OT neurons began to fire more to the newly-rewarded odor and less to the previously-rewarded odor, thus demonstrating that OT neurons are sensitive to the valence of odors. A provocative interpretation of these results, in conjunction with the known anatomy of the OT, is that this area is involved in learning associations between odors and reward. As a part of the striatum, however, action selection (i.e. licking for the water reward) as well as the actual receipt of reward could explain, at least in part, responses to rewarded odors and thus confound the
interpretation of the OT’s role in representing the valence of olfactory stimuli. To address this issue, the same group developed a new task that required mice to initiate and maintain licking for two seconds in order to receive a water reward, in the absence of any odor cues [Gadziola 2016]. OT neurons fired both at the initiation of licking and at the time of reward delivery. Therefore, the precise role of OT in learning associations between specific odors, actions, and outcomes has not been resolved. This topic is a major focus of the work presented in this thesis.

Attempts to assess the contribution of specific olfactory brain regions to the learning of odor-reward associations by lesioning has not revealed a single necessary region. Large lesions of the ventral brain that spanned many olfactory regions did impair the ability to learn odor-reward associations [Slotnick 1980; Slotnick 1985]. These lesions also typically included severe damage to the LOT, thus depriving areas that were not directly lesioned of olfactory sensory input. However, lesions of the LOT that failed to deprive sensory input to all brain regions, or local lesions encompassing only one or a few olfactory areas, did not impair the animal's ability to discriminate odors. The ability to compensate for focal lesions is potentially explained by the redundancy provided by the highly parallel output of the OB and the minimal requirement that multiple olfactory regions participate in at least one behavior that relies on the capacity to segment the space of potential odors to some degree.

The necessity of individual olfactory brain regions becomes apparent in tasks that place demands on short-term or long-term memory. For example, piriform lesions impaired the ability to perform an odor delayed non-match-to-sample task despite an intact capacity to learn and execute a go/no-go odor-reward association task [Zhang 1998]. Similarly, lesions of the horizontal nucleus of the diagonal band (HDB) impaired the ability of rats to execute the correct go/no-go action when a delay between odor and response period greater than 15 seconds was enforced, but not at shorter intervals [Roman 1993]. The HDB, like LEC, provides olfactory input to the hippocampus, which might account for the impairment arising only once demands on
memory were imposed. A complete understanding of the contributions to behavior of all olfactory regions will likely require many more such studies that utilize tasks with distinct cognitive demands.

1.8. **Odor-reward coding in high-level brain regions**

Orbitofrontal cortex (OFC) and basolateral amygdala (BLA) are heavily and reciprocally connected to the olfactory system, primarily the piriform cortex. Neurons in the OFC are sensitive to both odor identity and valence, and similar proportions of neurons respond to odors with positive and negative valence [Schoenbaum 1998; Schoenbaum 1999; Ramus 2000; Roesch 2007]. Unlike the OFC neurons, most BLA neurons respond to odors associated with a negative valence whereas relatively few respond to odors associated with a positive valence [Schoenbaum 1998; Schoenbaum 1999]. In odor-reward contingency reversal experiments, more than two-thirds of OFC neurons and all BLA neurons changed their odor responses [Schoenbaum 1999; Roesch 2007], demonstrating that valence coding took priority over odor identity coding in the OFC and BLA. During initial odor-reward contingency learning, as well as contingency reversals, BLA responses changed prior to the animal reaching behavioral criterion whereas OFC responses changed after criterion was reached [Schoenbaum 1999]. Based on this result, the authors proposed a two-stage model, in which the BLA is responsible for learning odor-reward associations and OFC is responsible for task execution based on the learned associations. In support of this model, 1) sensory inputs to BLA undergo plasticity during learning of cue-reward associations [Tye 2008] and 2) BLA lesions disrupt sensory-cued reward-motivated behaviors [Hatfield 1996]. Interestingly, the primary target of the BLA within the olfactory system is the pPC [Johnson 2000; Majak 2004], which has been proposed as an explanation for the presence of valence-related activity in pPC but not aPC [Calu 2007]. The
contributions of non-olfactory brain regions to reward-related responses in the olfactory system require further investigation.

1.9. Summary

Rodents are highly adept at learning arbitrary associations between odors and reward. The olfactory system is highly interconnected with the reward system as well as cognitive brain regions involved in reward-motivated behaviors. The pPC and OT are situated at the intersection of sensory and reward processing, and recent studies find that reward has a substantially greater influence on the activity in these areas than in areas immediately preceding them in the feedforward sensory processing hierarchy. These two areas are themselves hierarchically arranged, with the pPC providing feedforward input to the OT. However, little is known about the representation of odors in each area, how odor representations transform between pPC and OT, and the manner in which reward associations contribute to those representations. The work described in this thesis aims to shed light on these issues.
2. Experimental Methodology

2.1. Experimental animals

All experimental animals were C57Bl6/J mice acquired from Charles River Laboratories, aged two to four months at the start of experiments. Following implantation of a tetrode drive, nasal cannula, and head plate, all mice were housed individually. Behavior and physiology experiments took place over the course of one to two months. After the completion of all experiments, mice were euthanized and post-mortem histology was performed to confirm the location of electrophysiological recordings.

2.2. Tetrode and nasal cannula implantation surgery

Surgery was performed on naive animals and all behavioral training began after recovery from surgery. Mice were anesthetized with an IP injection of a mixture of Xylazine (10 mg/kg) and Ketamine (80 mg/kg). A craniotomy (~1mm) was made over the nasal canal (2mm anterior of nasal/frontal fissure, 1 mm lateral) on the skull, with a goal of implanting a nasal cannula to monitor sniffing during behavioral and physiological experiments. An 18 gauge stainless steel cannula (~5mm in length) was inserted into the craniotomy. To affix the cannula to the skull, superglue was used for initial placement followed by two applications of dental cement to ensure stability of the cannula. Additionally, a craniotomy (~1mm) was made over the dorsal skull at a location directly above the areas targeted for electrophysiological recordings with the goal of implanting a tetrode bundle (6 tetrodes plus a 200um-diameter optic fiber to ensure stability). The target locations were: pPC (coordinates: 0.5mm posterior, 3.8mm lateral, and 3.8mm ventral from bregma) and OT (coordinates: 1.2mm anterior, 1.5-2mm lateral, and 4.6mm...
ventral from bregma). To ensure stability of the animal’s head during behavior and recording at a later stage, a custom-made head plate (made of light-weight titanium, dimensions of 30 x 10 x 1 mm and weight of 0.8 g) was affixed to the skull. A shallow well was drilled over the posterior lateral skull and a single skull screw was affixed at that location. A wire was attached to this skull for grounding electrophysiological recordings. At the end of the surgery, a removable plastic cap was placed on the top of the nasal cannula to prevent foreign objects from entering the cannula. In addition, a plastic cone was positioned around the tetrode drive, and capped with a removable lid, to prevent damage to the drive. Following the completion of the surgery, mice were given one week to recover.

2.3. **Odor stimulus delivery**

A custom-made 16-odor olfactometer was used to deliver stimuli during behavioral tasks [Walker 1986]. The olfactometer had a carrier stream of air calibrated to flow at 1 liter per minute. A single odor at a time could be added onto this carrier stream by opening an odor-specific valve to permit airflow (at 0.1 liters per minute) from an input manifold, through a tube containing liquid-phase odorant, and finally into the carrier stream. One-way check valves prevented flow of odor from odorant-containing tubes into the carrier stream or back into the input manifold when air was not actively being flowed. Odor delivery to the animal was gated by a single (“final”) valve that directed odor to the animal or, between odor presentations, to an exhaust system. When odor was not flowed to the animal, a stream of clean air of the same flow-rate was directed to the animal. This ensured active clearing of odors between trials. In addition, an exhaust system cleared odors from the behavior chamber. To ensure immediate delivery of odors, and comparable odor concentration from trial-to-trial, the line to the final valve was primed with odor from the current trial beginning at the end of the previous trial (i.e.
minimum of 5 seconds, sufficient to replace the line volume 3-5 times). Sound from the switching of an odor-specific valve to perform selection of an odor for the upcoming trial was masked by a substantially louder non-odor-specific valve turned on simultaneously. The valves corresponding to each odor were randomly switched occasionally between sessions to prevent animals from learning sound, rather than odor, cues. Finally, all tubing that was odor-specific was replaced at the same time as new odors were added to the olfactometer in order to prevent odor from previous sessions lingering and being used as cues for task performance.

2.4. Respiration recordings

Respiration was recorded through a nasal cannula implanted during surgery. Following surgery, the nasal cannula was cleaned daily to prevent clogging. In this way, nasal cannulae typically stayed clear continuously for months of recordings. During experiments, a plastic tube connected to a pressure sensor was fitted over the nasal cannula to form a continuous pressure environment. Pressure signals were amplified 10X, bandpass filtered 0.1-100Hz, and recorded at 1000 samples per second. For analysis, the start of each inhalation was identified as negative-going zero crossings in the pressure signal. Likewise, the end of each inhalation (i.e. start of exhalation) was identified as a subsequent positive-going zero crossing in the pressure signal.

2.5. Electrophysiological recordings and analysis

Neural activity was recorded with drives containing six tetrodes [Gray 1995]. The tip of each wire was gold-plated until the impedance was 250-450 kOhms. Electrophysiological signals were acquired and amplified with a custom-made system built on two 16-channel analog chips designed by Intan Technologies. Each channel was digitized, sampled at 20kHz, and
bandpass filtered with a 2nd-order Butterworth filter 500-3000Hz. Potential spike events were detected as any activity that crossed 3.7 standard deviations of noise (corresponding to 1:10,000 events by chance). Potential spike events were then manually clustered with Mclust software [Redish 2008] to identify single-unit activity. The quality of clustering was ensured by requiring than 1 in 1000 spikes to have occurred within a 2ms refractory period of one another, clear overlap of single spike waveforms, and L-distance below 0.05. Following the completion of all behavior and recordings, tetrode placement was confirmed with an electrolytic lesion and postmortem histology. To minimize selection bias, all spikes that were visible during each session were recorded, sorted and analyzed. At the end of the recording session each day, the entire bundle of tetrodes was lowered 40um to obtain a new set of neurons for the subsequent day.

2.6. Behavioral training

When animals have fully recovered from surgery (approximately 7 days), they were placed on a water deprivation schedule for behavioral conditioning. Animals were first acclimated to head restraint (with their implanted head plate held in place) on the behavior rig over the course of 3-5 days. Then, animals were trained to lick for water from a spout positions 2-3 millimeters from the mouth until satiated (typically 1-1.5mL); licking this spout (or not licking) would subsequently serve as the response for all behavioral tasks (go/no-go). Lick training took approximately 2-4 additional days. Next, odors were presented for one second with a varying (5 second mean) inter-stimulus interval, and reward was only available for a subset (half) of these odors. At first, only two odors (one rewarded, one unrewarded) were presented in order to expedite learning. The number of odors was increased to four and then eight (i.e. full task odor panel). Initially, water reward would be available immediately upon licking during odor
presentation. Gradually, the delay between lick/odor presentation and reward delivery was increased until reward delivery occurred at least 500 milliseconds after the conclusion of odor presentation. Neural recordings began once the animals reach strong performance (>90% correct) on the full eight-odor panel. No odors were replaced during this initial period in order to facilitate the acquisition of task structure. In the final stage of experiments, a subset of odors were replaced between some sessions in order to study learning of novel odor-reward associations, as described in the main text.
3. **Odor and Reward Coding in pPC and OT**

Sensorimotor decision making is the process of selecting and executing an appropriate motor action in response to a given sensory stimulus. Rodents can successfully learn to associate odor cues with motor responses following only a few repetitions of trial-and-error per odor when the correct response is reinforced by reward [Schoenbaum 1995; Slotnick 2000; Doucette 2008; Gire 2013; Gadziola 2015]. The resulting decision process can be highly odor-specific: mice can accurately determine the presence or absence of a single rewarded odor hidden within previously unexperienced mixtures of more than a dozen distractors with similar molecular features as the target rewarded odor [Rokni 2014]. Although rodents make strong and flexible behavioral associations between odors and reward, the interaction of sensory and reward coding in the olfactory system during decision-making is poorly understood.

The rodent olfactory system is arranged predominantly in a feed-forward hierarchy that ultimately converges on two high-level sensory areas, the posterior piriform cortex (pPC) followed by the olfactory tubercle (OT) [Mori and Sakano 2012]. pPC is an association cortex that is heavily interconnected with multiple brain regions implicated in reward-related behaviors, including orbitofrontal cortex, amygdala, and the midbrain dopaminergic system [Haberly 1978; Lushkin 1983; Datiche 1996b; Johnson 2000; Haberly 2001]. OT is the olfactory striatum, consisting of primarily (>90%) D1- or D2-receptor-expressing medium spiny neurons and having similar inputs and outputs as the dorsal striatum [Wesson 2011; Giessel and Datta 2014]. Unlike dorsal striatum, OT is specialized for olfaction with a well-defined cell body layer that receives input directly from OB in addition to input from multiple olfactory cortical areas including pPC [Scott 1980; Schwob and Price 1984; Sosulski 2012]. In turn, OT is by far the largest source of olfactory input to midbrain dopaminergic neurons [Watabe-Uchida 2012], suggesting a
prominent role in reward-related behaviors.

Activity of single-neurons that correlate with choice and reward has been observed in both pPC and OT [Calu 2007; Gadziola 2015]. The anterior piriform cortex (aPC), which immediately precedes pPC and OT in the anatomical hierarchy of olfactory regions, does not have choice activity in the first sniff of an odor stimulus [Miura 2012], when decision making occurs [Uchida and Mainen 2003; Rinberg 2006], although post-decision activity in aPC is selective for reward [Gire 2013]. Single-neuron reward-selectivity emerges prior to the motor response at the level of OT neurons [Gadziola 2015], however it is not known whether reward-selectivity is present at the same time in pPC as well. If reward-selectivity is available in pPC during decision making, downstream OT could inherit this selectivity. Alternatively, heavy interconnection with the reward system provides the OT with the necessary reinforcement signals and plasticity-inducing dopamine for learning to represent arbitrary odor-reward associations [Wieland 2015; Gadziola 2015].

Although reward-related activity is present in pPC and OT, the relationship between stimulus coding and reward is not known [Calu 2007; Gadziola 2015]. For instance, it is not clear whether odor-selectivity and reward-selectivity are encoded by distinct or overlapping neuronal populations in these areas. To address this issue, we trained mice on an odor-reward categorization task in order to probe stimulus-tuning simultaneously with reward-selectivity. We recorded the activity of single neurons in the pPC and OT as mice learned, through trail-and-error, the reward valence assigned to a panel of previously unexperienced odor stimuli. This approach allowed us to study and dissect the interaction of odor coding with reward coding at two crucial junctions of the rodent olfactory system.

Behavioral responses confirmed that mice reliably learned odor-reward contingencies within a single session, permitting the evolution of neural activity to be monitored throughout the learning process. Along with neural recordings, we simultaneously measured sniffing in order to
reveal the earliest components of sensory processing and decision making during the first sniff of odor cues. We observed a transformation from a code in the pPC that was dominated by odor selectivity but not reward selectivity to a code in the OT that explicitly represented reward with a loss of odor selectivity. Furthermore, pre-decision reward selectivity emerged in the OT in tandem with learning during the first session of experience of novel odor-reward associations. Together, these results suggest that reward information is not explicitly available in olfactory cortex during decision-making, rather reward associations are computed outside of sensory cortex in circuits involving, but not necessarily limited to, olfactory striatum.

3.1. Behavioral task

Go/no-go is a robust, operant conditioning paradigm that can be used to teach animals novel stimulus-reward associations and measure learning directly via the behavioral response [Schoenbaum and Eichenbaum 1995; Bodyak and Slotnick 1999; Calu 2007]. In short, a stimulus is presented and the animal must decide, based on the identity of the stimulus, whether to respond (Go) or not respond (No-Go). For all of the experiments described below, water-deprived mice learned to discriminate odors that indicated the availability of a water reward (referred to as rewarded odors) from odors that indicated the absence of a water reward (referred to as unrewarded odors). A panel of eight odors, half of which were rewarded and half unrewarded, were presented on randomly interleaved trials during each experimental session. This size panel was chosen as a compromise between the competing goals of (1) having as many odors as possible to characterize the sparseness of odor tuning and (2) having as many trials as possible per odor to accurately measure behavioral and neural responses to the odor.

During the task, mice were held in place on a behavior station with a head plate that was surgically affixed to the skull (Figure 1a). The heads of the mice were fixed in place to
Figure 1: Behavioral task design. a) A mouse in the head-fixed behavior setup. The odor port (red dot) and lick port (yellow dot) are labeled. b) Diagram of the task structure, indicating the periods for odor, response (lick/no-lick), and reward delivery.
odors directly to their nose with precise control over stimulus timing and concentration and (2) minimize the delay and variability of the motor behavior of the animal's Go response. After 5-7 days to recover from surgery, mice began water deprivation and trained for about 3 days to tolerate head restraint. Concurrently, mice learned that water was available ad libitum during head restraint from a water spout that was positioned in front of the mouth (maximum one 3-5uL drop per second). Odor training began once mice licked for a minimum of 150 drops in a single session. Odors were delivered from a custom-built olfactometer [Bodyak and Slotnick 1999; Walker and O'Connell 1986] through a port in a fixed position in front of the nose, resulting in reproducible odor timing and concentration. Every few seconds (5s mean, exponentially-distributed inter-stimulus interval) one odor was randomly chosen from the panel of eight odors and was presented for one second (Figure 1b). When odor was not being presented between trials, clean air was continuously flowed over the nose of the mouse and removed with an exhaust flow in order to minimize cross-contamination of odors from trial-to-trial. In the early phase of training, licking was rewarded with water delivery while the rewarded odor was still flowing to the animal in order to make clear the association among odor presentation, licking, and reward delivery. Reward delivery was progressively postponed over the course of 2-3 sessions, slow enough to avoid a setback in behavioral performance. Ultimately, reward was delivered a variable 550-650ms after the one second odor presentation had concluded in order to distinguish effects related to reward delivery from sensory processing in neural recordings.
Figure 2: Task performance. **a)** Average task performance over the first four sessions for 14 mice. The hit rate, correct rejection (CR) rate, and overall performance rate are shown separately. **b)** Lick rate by trial type averaged over sessions in well-trained mice (i.e. after session 5). Shading indicates standard error of the mean. **c)** Raster plots of individual lick times by trial, shown separately for rewarded and unrewarded odors.
Figure 3: Introduction of novel odors to the stimulus panel. During the initial phase of training on the task structure, the eight odors in the stimulus panel remained the same. That is, odors were not replaced between sessions. After 5-10 sessions, in which time the mice learned to perform the task successfully with this initial odor panel, half of the odors (2 rewarded and 2 unrewarded) were replaced between some sessions to study learning of novel odor-reward associations (“Novel Sessions”). Before changing the stimulus panel again, the stimulus panel was held constant for at least one additional session in order to measure behavioral and neural responses to the newly-familiar odors. Therefore, sessions with all familiar odors and sessions with novel odors were intermingled. Each of the recorded neurons was only recorded during a single session, and therefore had a different stimulus panel than neurons recorded during other sessions; the different colors of neurons in the figure indicate different neurons. Different shapes indicate different sets of odors, with odors represented by the same shape being introduced at the same time. Hollow shapes indicate novel odors that the mouse has not been exposed to in previous sessions. Filled shapes indicate odors that have been learned in a previous session (“Familiar odors”). Red shapes indicate rewarded odors and black shapes indicate unrewarded odors.
Figure 3 (Continued)

<table>
<thead>
<tr>
<th>Initial Task Learning</th>
<th>Rewarded Odors</th>
<th>Unrewarded Odors</th>
<th>Recorded Neurons</th>
</tr>
</thead>
<tbody>
<tr>
<td>Day 1:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Day 2:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Day 3:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Day 4:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Day 5:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Day 6:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Day 7:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Day 8:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Day 9:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Day 10:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Day 11:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Day 12:</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Mice acquired the task structure and reached a high level of performance (>90% correct trials) within three to four sessions from the start of training with odors (one session per day, consecutive days; Figure 2a). Among correct behavioral responses, the fraction of go responses on rewarded odor trials (i.e. hit rate) started higher and peaked earlier than the fraction of no-go responses on unrewarded odor trial (i.e. correct rejection rate). This behavioral bias towards licking likely reflects the asymmetry inherent to the go/no-go paradigm that only go trials had the possibility to be rewarded. In a typical session, mice worked for about 300-500 trials and collected 150-250 water rewards before they were satiated, which ensured at least 20-30 repetitions per odor. Well-trained mice consistently began licking within 200-400ms of odor onset (Figure 2b and 2c: top) for rewarded odors and maintained licking until reward delivery. In contrast, mice did not even initiate licking on most unrewarded odor trials (Figure 2c: bottom), indicating that the animals became confident of the reward association early during odor presentation.

To discourage a high false alarm rate, a hit was only scored on trials that the mouse licked in at least three out of four 350ms bins, beginning at the time of odor onset. It follows that mice were required to make a decision before the end of the second bin, 700ms into the odor presentation, but this did not seem to force early inaccurate decisions. Again, mice initiated licking on hit trials several hundred milliseconds before the end of the second bin and, on miss trials, mice rarely began licking around the end of the second bin. Overall, mice maintained a high hit rate and low false alarm rate. No punishment was given in the case of misses, beyond the loss of an opportunity to receive a water reward, or false alarms. Thus, the low frequency of licking on unrewarded odor trials suggests that the energetic expense of licking was a sufficient deterrent to employing an “always go” strategy to maximize reward. The rate of licking, on average, was twice as high on hit trials than false alarm trials (Figure 2b), indicating greater confidence in the reward association of rewarded odors. Interestingly, mice occasionally
began licking after the stimulus presentation concluded on miss trials. Similar behavior did not occur on correct rejection trials, demonstrating an analogous confidence in the reward association of unrewarded odors on correct trials.

Naive mice were initially trained on the task with the same panel of eight odors each day (one session per day) until they achieved successful performance (>90% correct) for two consecutive sessions (Figure 3). At first, only two of these eight odors (one rewarded, one unrewarded) were presented in order to expedite learning. The number of odors was increased to four and then eight in subsequent sessions as performance improved. This phase of learning the task structure took 5-10 sessions. Following successful learning of the task structure, novel odors that were not among the eight odors in the original panel were introduced in order to study learning of odor-reward associations. Four novel odors replaced four odors that had been used in a previous session (“Novel Session”). Specifically, two novel rewarded odors and two novel unrewarded odors replaced two familiar rewarded odors and two familiar unrewarded odors, respectively. The odor panel remained the same for one or two additional sessions in which behavioral and neural responses to the newly-familiar odors were measured. This process was repeated with daily sessions for one to two months until all recordings in the mouse were complete.

Ultimately, the data collected from each mouse comprised a large pool of odor-reward associations, aggregated from individual sessions involving only a small portion of the entire pool; the stimulus set for any given session only partially overlapped with a minority of other sessions. In our efforts to study reward coding, this strategy helped to minimize the contribution of stimulus tuning for any single odor to the overall ability to successfully categorize rewarded versus unrewarded odors. Sections 3.3 to 3.7 include analysis of coding across this large pool of odors in sessions after successful learning (“familiar odors”), then sections 3.8 and 3.9 describe the changes in behavioral and neuronal responses during learning (“novel odors”).
3.2. Sniff behavior

Sensory transduction in olfaction begins when odors reach olfactory receptors in the nasal epithelium located near the back of the nasal canal. Thus, the start of sensory processing on any given trial was not identical to the time of odor delivery from odor tubes to the animal's nose. Previous studies have reported a range of inter-sniff intervals for mice of approximately 80-500ms [Wesson 2008], corresponding to a rate of 2 to 12 sniffs per second. Since decisions are made in a few hundred milliseconds, the trial-to-trial variability in sniff timing could have hidden fine temporal structure in behavior or neural activity. In order to access this timing

Figure 4: Sniff behavior during task performance. a) An example trace of five seconds of respiration measurements obtained through air pressure from a nasal cannula. Exhalation is positive-going pressure and inhalation is negative-going pressure. b) Average sniff rate over all trials from an example behavior session. c) Histograms of the times of the first, second, third, and fourth sniffs of each trial.
information, airflow was measured through a cannula implanted in the nasal canal (Figure 4a). The nasal cannula and tetrode drive were implanted during the same surgery, but always on opposite (i.e. contralateral) sides. Each lateral olfactory tract delivers sensory input from a single olfactory bulb to the ipsilateral pPC and OT. Implanting the nasal cannula and tetrode drive on opposite sides was intended to avoid altering the natural airflow to the relevant sensory neurons for our recordings.

The rate of sniffing in our experiments was consistent with previous findings in mice [Wesson 2008]. Once familiar with the task structure, the rate of sniffing during odor presentation did not differ from the spontaneous rate between trials (Figure 4b). As a result, there were typically 3-5 sniffs for each one second odor presentation (Figure 4c). Immediately following odor onset, however, there was often a brief (~100ms) increase in sniffing. This short period was comparable to the minimum duration of a single sniff, and corresponded to an extra sniff taken at the beginning of odor sampling on each trial. The ability of the mice to quickly react to odor presentation was not surprising because a click of the valve that controlled delivery of odorized air was audible from within the behavior chamber. This single valve was common to all odors and thus was not informative about odor identity.

3.3. Single unit responses of pPC and OT neurons

To isolate and record single units, multi-tetrode drives [Gray 1995] were implanted in either pPC (n=6 mice) or OT (n=8 mice) during a stereotactic surgery. Following the completion of all behavior and recordings, tetrode placement was confirmed with an electrolytic lesion and postmortem histology (Figure 5). Peri-stimulus time histograms (PSTHs) are shown for representative pPC and OT cells in Figure 6. Since odor responses began following the first sniff during odor presentation, all of the remaining analyses were performed using the time of
the first sniff as the start time for each trial. An example of sniff-aligned and odor onset-aligned activity for the same OT cell is shown in (Figure 6a) for comparison. Figure 7a-d shows responses for all significant cell-odor pairs in both areas with rewarded and unrewarded odors separated. Rather than absolute firing rate, the plots show the area under the receiver-operating characteristic curve (auROC) for each cell-odor pair relative to baseline (i.e. pre-stimulus) firing rate. This auROC gives a measure of how well the firing rate at any given time can be discriminated from the baseline firing rate for that cell. Briefly, the value of the auROC for a given time bin indicates the percentile of that time bin in the distribution of baseline firing rates for bins of the same width. Therefore, a firing rate that is exactly the median of the baseline firing rate distribution would have an auROC of 0.5. Excitatory responses (yellow) correspond to an auROC greater than 0.5 but less than or equal to one, whereas inhibitory responses (blue) correspond to an auROC less than 0.5 but greater than or equal to zero. For analysis as well as data visualization, the auROC provides a more consistent measure to compare across neurons than absolute firing rate.

Baseline firing of pPC neurons was invariably low (Figure 8a). Odor responses were
Figure 6: Single-unit responses of pPC and OT neurons. a) An example OT cell. Spike rasters are shown unaligned (top) and aligned (middle) to the first inhalation of odor. A peri-stimulus time histogram (PSTH) aligned to the first sniff is shown directly below the rasters. Four PSTHs for the rewarded odors are shown in red and four PSTHs for the unrewarded odors are shown in black. b,c) PSTHs for example cells with odor responses that are sustained throughout the one-second odor delivery. d,e,f) PSTHs and spike rasters for example cells with transient odor responses during the first sniff of an odor. b,d were recorded in pPC and a,c,e,f were recorded in OT. All PSTHs shown are aligned to the beginning of the first inhalation during the odor presentation.
Figure 7: Population odor responses in pPC and OT. a-d) Area under the receiver-operating characteristic (auROC) for each cell-odor pair. For ease of comparison, cells recorded in pPC and OT are shown separately, as are rewarded and unrewarded odors. Time of zero is odor onset. Yellow indicates that the cell was excited relative to its baseline firing rate, whereas blue indicates that the cell was inhibited relative to its baseline firing rate. For visualization purposes, cell-odor pairs with noisy firing during pre-odor time are not shown, resulting in a different number of cell-odor pairs between rewarded and unrewarded odors despite and equal number of recorded cell-odor pairs.
Figure 7 (Continued)

a) pPC rewarded

b) OT rewarded

c) pPC unrewarded

d) OT unrewarded
predominantly excitatory (Figure 7a,c). As expected, neurons with excitatory responses showed odor selectivity: these neurons had differential odor responses and often responded to only a subset of odors. Most cells continued to respond for the duration of the one second odor presentation (Figure 6b and 7a,c). Other cells had a strong transient (100-200ms) response well aligned to (and presumably driven by) the first inhalation of an odor, but did not respond for the remainder of the odor presentation despite continued sniffing (Figure 6d). Interestingly, the cells that were inhibited appear to be inhibited by all odors non-selectively. Low baseline firing rates made it difficult to detect odor-selective inhibition, if it were present, and suggest that these neurons did not directly contribute to odor discrimination. Therefore, inhibitory responses of pPC neurons were not analyzed further.

The responses of OT neurons were clearly distinguishable from those of pPC neurons in several ways. Higher baseline firing rates (e.g. greater than 5 spikes/s) were much more common in OT (Figure 8b). Unlike pPC, a substantial portion of neurons were inhibited during odor presentation, often following a brief period of excitation (Figure 7b,d: bottom). Transient excitatory responses at the beginning of odor presentation were much more prevalent in the OT.

Figure 8: Single-unit firing properties in pPC and OT. a,b) Baseline firing rates for all cells recorded in pPC and OT, respectively. Note: the last bin in b shows cells that had a firing rate greater than or equal to 60 spikes per second. c) Lifetime sparseness for cells recorded during sessions in which the animal was familiar with all eight odors in the stimulus panel.
(Figure 7b,d). Notably, very few responses to unrewarded odors lasted more than 500ms. Excitatory and inhibitory responses lasting longer than 500ms were more common in OT for rewarded odors than unrewarded odors (Figure 7b vs Figure 7d). In contrast, the overall distribution of response profiles did not differ between rewarded and unrewarded odors for pPC neurons (Figure 7a vs Figure 7c).

The strength of odor selectivity, or tuning, of individual cells can be quantified in terms of lifetime sparseness [Wilmore 2001; Bhandawat 2007]. Briefly, we used lifetime sparseness to measure how different were the average responses of a given cell to each of the eight odors.

\[
\text{Lifetime sparseness} = \frac{8}{7} \left( 1 - \frac{\sum_{j=1}^{8} |r_j|/8}{\sum_{j=1}^{8} r_j^2/8} \right)^2
\]

where \( r_j \) is the (baseline subtracted) firing rate response of the cell to odor \( j \) during the first 200ms following the start of odor sampling. Lifetime sparseness ranges from 0 to 1, with zero corresponding to identical responses to all eight odors and one corresponding to a perfectly selective response to a single odor. The distributions of lifetime sparseness for pPC neurons and OT neurons were overlapping but distinct. The peak of the OT distribution was near zero and therefore OT was dominated by cells with non-selective odor responses (Figure 8c: blue). In contrast, lifetime sparseness of pPC cells was distributed fairly evenly throughout the range from zero to one (Figure 8c: yellow). Though distinct, the populations of neurons in both pPC and OT spanned a wide range of odor selectivity.

3.4. Influence of sniffing on odor responses

Sniffing behavior imposes temporal structure on the delivery of sensory input to the olfactory system. We investigated the extent to which sniffing shaped the odor responses of
neurons in pPC and OT. To quantify sniff modulation of odor responses, we calculated an index that measures how spikes were distributed throughout the phases of the sniff inhalation-exhalation cycle. Every sniff cycle was split into eight bins that each swept through $\pi/4$ of the complete sniff, and spikes were assigned to the corresponding bin. Then a vector sum was calculated for each cell of the average firing rate across the eight sniff-phase bins (e.g. Figure 9a,b); only sniffs that occurred during odor presentation were included in this analysis. Finally, the vector sum was normalized to the total firing rate across all bins. The resulting modulation index ranged from 0 (equal firing rate across all phases of the sniff) to 1 (all firing during a single phase bin), and could be compared across cells. Furthermore, the inhalation-exhalation cycle of sniffing was not perfectly periodic, rather mice could extend or hasten the time before the next inhalation to modulate the inter-sniff interval. Spikes that occurred during this inter-sniff interval would have been concentrated in the one or two phase bins at the end of exhalation, before the next inhalation began. Thus, our single index was designed to capture the sniff-modulation of spikes across different phases of a rapid sniff as well as differential spiking in the sniff versus inter-sniff interval.

Although pPC and OT both contained cells spanning the full range of possible modulation indices (i.e. 0 to 1), highly sniff-modulated cells were common in pPC but rare in OT (Figure 9c vs Figure 9e). To eliminate the influence of odor tuning on the measure of sniff modulation, the data were analyzed for each cell-odor pair separately. Potentially, a cell could have had a large sniff modulation index due to the chance occurrence of spikes in one or a few phase bins. Such an effect could have resulted from cells with low firing rates, very few sniffs per trial, and/or few trials per odor. To determine the statistical significance of a given modulation index, we compared the calculated index with indices calculated from shuffled data. Specifically, we shuffled the sniff data across trials with the same odor during the same session. In this way, the overall statistics of sniffing were preserved but the relationship between sniffing
Figure 9: Influence of sniffing on odor responses. a) Peri-stimulus time histograms of the odor responses of a cell that was moderately modulated by sniffing during the odor presentation. The response was aligned to the first inhalation after the odor onset (t=0). b) Representation of the vector-sum method used to calculate the sniff modulation index for the cell in a. Black arrows illustrate each of the eight bins of equal phase coverage through the sniff cycle, with magnitude equal to the firing rate in the bin normalized to the total firing rate throughout all bins. c,e) Histograms of sniff modulation indices with significant indices indicated in red. d,f) Cumulative distribution of p-values for sniff modulation indices. g) Habituation of odor responses across sniffs during the same odor presentation. h) PSTH for the odor response indicated by the arrow in g. i,j,k) Odor responses for the cell in h for the first, second and third sniffs of the odor presentation.
Figure 9 (Continued)
and neural activity was removed. We repeated this process for 500 shuffles and calculated a p-value for each cell that gave the probability of obtaining a modulation index with shuffled data greater than or equal to the true modulation index. The cell-odor pairs with a p-value below 0.05 are indicated in red in (Figure 9c,e). Further illustrating the extent of sniff modulation in pPC, significant modulation indices were concentrated near one.

Rather than rely on a single threshold (such as p<0.05) to determine whether activity in the pPC and OT was genuinely modulated by sniffing, we sought a more sensitive and robust measure of significance at the population level. To this end, we plotted the cumulative distribution of p-values generated in the shuffle analysis for each area (Figure 9d,f). If odor responses were not affected by sniffing, the data should have fallen along the unity line (i.e. dashed line in the Figure 9d,f). Formally, such a distribution would be equivalent to selecting a modulation index generated from the shuffled data at random and treating that index as the true modulation index. The actual distributions for pPC and OT dip below the unity line, indicating that the distributions of p-values were skewed towards zero (i.e. more significant sniff modulation) more than would have been expected by chance. In other words, pPC and OT activity were both significantly modulated by sniffing at the population level, though pPC had substantially more sniff modulation that was apparent at the single-cell level.

To further explore the sniff modulation of pPC neurons, we characterized the sniff-to-sniff change in response through the course of an odor presentation. Aligning the spikes of pPC cells to their respective sniffs during odor presentation revealed rapid habituation. On average, the odor response of pPC cells was negligible by the third sniff (Figure 9g). Although some pPC cells only responded to odors during the first sniff (e.g. Figure 6d), other cells did continue to respond for the remainder of the odor presentation (e.g. Figure 6b). In the rare case of a sustained response, nearly all firing above baseline during the odor presentation was well aligned to inhalations (Figure 9h-k). These analyses revealed that sniffing imposes a temporal
structure on the odor responses of pPC neurons that had not previously been observed.

3.5. **Reward coding at the single-cell level in pPC and OT**

Next, we wanted to characterize reward coding in the pPC and OT. In this section, and the following two sections, analysis was restricted to familiar odors (i.e. odors that the mouse successfully learned in a previous session). The last two sections of this chapter describe the evolution of responses to novel odors. To start, we examined the extent to which reward information was explicitly available in the firing rate of individual neurons. Specifically, we asked how well the firing rate of a given cell could discriminate rewarded odor versus unrewarded odor trials. This was calculated as the auROC when comparing the distribution of responses to all rewarded odors and the distribution of responses to all unrewarded odors. The analysis revealed a striking contrast in the selectivity for reward valence between pPC and OT neurons (Figure 10a,b). Valence selectivity was substantially more prevalent in OT than pPC both during and after odor presentation. Furthermore, selectivity for valence was achieved with higher firing rates for rewarded odors than unrewarded odors (yellow in Figure 10a,b) as well as higher firing rates for unrewarded odors than rewarded odors (blue in Figure 10a,b). In fact, some cells switched preference (i.e. higher firing rate) between rewarded and unrewarded odors during the course of the response.

We then asked whether valence selectivity emerged at times corresponding to particular task events, such as stimulus onset or reward delivery. In other words, we wanted to know the time at which valence selectivity first became significant for each cell. In order to measure the significance of valence selectivity, we compared the auROC at times after odor onset against the distribution of auROCs during pre-stimulus time. Consistent with a lack of actual valence selectivity for pPC cells, the times at which the valence auROC first reached a liberal threshold.
Figure 10: Reward valence selectivity during decision-making. a, b) Area under the receiver-operating characteristic (auROC) to illustrate cells’ ability to discriminate rewarded versus unrewarded odors that had been successfully learned in previous sessions (i.e. familiar odors). Responses are aligned to the first inhalation following odor onset (t=0). The cells (rows) are hierarchically clustered by the first three principal components calculated from the initial (first 200ms) firing rate, delayed response rate (200ms-800ms), and response rate at the time of reward delivery (1500-2000ms). c) Histograms of the first time bin in which the auROC for a cell was significant for cells in pPC and OT. d) Zoom-in of the histograms in c, along with an overlay (hollow black bars) of the mean time of the first lick on hit trials to illustrate the beginning of the motor response.
Figure 10 (Continued)
of significance (p<0.01, with 35 comparisons per cell for all time bins) were scattered throughout the trial and not well-aligned to any task events (Figure 10c: orange). In contrast, valence selectivity emerged in 46% of OT cells within the first 300ms of odor sampling (Figure 10c: blue). The absence of any subsequent peaks indicates that there was not a separate population of OT neurons that responded only to later task events, such as reward delivery. Most of these selective responses emerged before the mouse began to lick on rewarded odor trials (Figure 10d). This motivated us to further examine reward selectivity in the pre-decision activity of pPC and OT cells.

Figure 11a illustrates a major obstacle to studying reward selectivity in sensory areas where the cells exhibit strong stimulus tuning, like pPC and OT. Cells that are highly selective for one or a few stimuli necessarily have both high lifetime sparseness and high reward valence selectivity. This is because the cell is selective for an odor, and that odor has a reward valence associated with it, so the cell’s activity technically discriminates both odor and reward. However, this is a trivial form of reward selectivity and does not necessarily imply that reward is actively or explicitly coded by the cell. At the other extreme, cells that respond indiscriminately to most or all odors (i.e. high lifetime sparseness) cannot be very selective for reward valence. Indeed, the pre-decision (first 200ms of odor sampling) responses in pPC and OT both follow this trend, and cells in the two areas that have similar lifetime sparseness also have similar average reward selectivity (Figure 11a).

If the cells in both areas do not explicitly code for reward, then the relationship between stimulus tuning and the resulting ability to discriminate reward valence should be random. To test this assumption, we shuffled the valence labels for each odor and re-calculated the reward valence auROC. We then repeated this procedure 100 times and compared the actual valence auROC of the cell with the distribution of shuffle-generated auROCs. This allowed us to ask how extreme (i.e. far from auROC=0.5) was the actual auROC of each cell relative to chance.
Figure 11: Influence of odor sparseness on reward selectivity. a) Mean reward valence selectivity, as measured by auROC, as a function of lifetime sparseness. auROC was calculated for the first 200ms from the start of odor sampling (first inhalation following odor onset). b) The percentile of the true valence auROC (from a non-discriminating value of 0.5) among the distribution of auROCs generated from shuffling the valence labels of the odors (lines show the median). Dots are individual cells. The dashed line indicates chance; that is, the true auROC falling at the median of the distribution on average. Error bars are 95% confidence interval on the median. The number of cells used to calculate each point in the plots are indicated in parentheses.
Specifically, we plotted the percentile of the auROC among the shuffle-generated distribution (Figure 11b). In the absence of true reward selectivity, the actual data should be near the 50th percentile on average. Indeed, pPC cells were not, on average, more extreme than the 50th percentile regardless of how narrow or broadly tuned were the cells (lifetime sparseness). On the other hand, OT cells were substantially more extreme than expected by chance. In fact, there was a trend for the valence auROC to be a larger outlier the more sparsely the cell responded to odors (Figure 11b: blue). In other words, more sparse odor tuning was related to an increase in the actual sensitivity to reward valence. Therefore, a substantial population of OT cells exhibited valence selectivity at the single-cell level in the few hundred milliseconds between odor onset and the animal’s motor response.

3.6. **Reward coding at the population level in pPC and OT**

Collectively, many individual cells with weak reward valence selectivity could still give rise to reliable valence selectivity at the population level. To determine whether our pPC and OT datasets had this property, we attempted to decode the valence of odors from single trial responses of cells in each area. An explicit code for valence at the population-level, by definition, should be possible to read out with a linear classifier [Pagan 2013]. For this reason, we chose to use a support vector machine with a linear kernel as the decoder. For each cell, the trials during the session it was recorded were randomly assigned to training (80%), validation (10%), and test (10%) sets. Since only a few cells could be recorded per session, cells were pooled across sessions and animals for this analysis. In this way, decoder performance was tested as a function of the number of cells used for classification.

Reward-related activity could have been present at any time during a trial following the beginning of odor presentation. Again, we focused specifically on the period prior to the initiation
of a motor response on rewarded odor trials in order to quantify reward information that might have been available during decision making. To empirically determine an appropriate time window to analyze, we decoded reward valence from OT activity for a series of windows beginning at the first inhalation of the stimulus (Figure 12a); specifically, we used the average firing rate response (i.e. baseline subtracted) for each cell within the window. Decoder performance increased with longer windows until saturating at 200ms. Accordingly, the remaining decoder analyses were performed using the average firing rate response within a 200ms window beginning at the first sniff of the odor.

A single neuron could respond differently to two odors, one rewarded and one unrewarded, because the response is valence-specific, but the differential response could just as plausibly be explained by odor tuning. In turn, the activity of a population of neurons with sparse responses to odors (such as pPC) could discriminate rewarded and unrewarded odors without requiring access to any information about the reward associations of those odors. If population activity explicitly codes for reward valence beyond sensory tuning alone, the ability to decode reward valence should not be sensitive to increasing the difficulty of decoding stimulus identity. Therefore, we tested the ability of our decoder to classify reward valence from progressively larger number of odors (e.g. 1 rewarded odor versus 1 unrewarded odor, 2 rewarded odors versus 2 unrewarded odors, etc.). This required the classifier to find a decision boundary that separated responses to an increasing number and diversity of odors. However, the size of the training set (i.e. number of single-trial responses) was held constant regardless of the number of odors used for classification to prevent classifiers trained on more odors from having better performance simply via access to more data. Indeed, decoder performance for the pPC data, which appeared to have weak single-cell valence selectivity at best, became substantially worse as the number of odors increased (Figure 12b). We compared this result to the decoding performance achieved by a classifier that was trained and tested on the same
Figure 12: Decoding reward valence from population data. 

**a)** Decoding accuracy for reward valence obtained with a support vector machine, trained and tested on the average response rate of OT cells in a window that began at the time of the first inhalation of an odor and had a duration equal to the time denoted for each curve. 

**b,c)** Decoding accuracy for similar classifiers that were trained and tested on the given number of odors, half of which were rewarded odors and half of which were unrewarded odors. 

**d,e)** The same as **b** and **c** but for classifiers trained and tested on odor responses with the reward valence label shuffled. 

**f,g)** The difference between the true decoding accuracy (**b,c**) and the decoding accuracy obtained on 10 shuffles of the data (**d,e**). Error bars in **a-e** show the standard error of the mean. Error bars in **f,g** show the standard deviation across shuffles to illustrate the spread of differences expected to arise from shuffle to shuffle.
Figure 12 (Continued)
dataset after shuffling the valence labels of the odors. This shuffling procedure would remove any potential explicit reward information and force the decoder to rely on stimulus-tuning alone for classification. Decoding performance on valence-shuffled pPC data followed the same trend as the actual pPC data (Figure 12d,f), suggesting the absence of an explicit code for valence in pPC. In contrast, decoder performance for the OT data did not decrease substantially as the number of odors increased (Figure 12c). On the other hand, performance did decrease for valence-shuffled OT data confirming that valence-specific information was lost in the shuffling process (Figure 12e,g).

The decoding analysis in Figure 12 pools over all odors and mice, so we sought to determine whether these results held across mice and specific sets of odors. The dataset for each mouse consisted of a variety of odors in order to study reward coding beyond the coding for odors alone (Figure 13). We then focused our analysis on sets of eight odors for which individual mice had at least a moderate number of cells (>5 cells) recorded during sessions in which all of those same odors were used. We reasoned that an explicit code for reward should have the same structure regardless of the particular stimuli that carry the reward association. Therefore, we trained linear classifiers (SVMs) to decode reward from the responses to a set of four odors, then tested the ability of the same classifier to decode reward from the responses to a different set of four odors (Figure 14). Overall, we found that classifiers trained on OT data successfully decoded reward from responses to odors that the classifiers did not see during training. On the other hand, classifiers trained on pPC data did not have this property. Importantly, these results were consistent across individual mice (OT: n=5 mice, pPC: n=3 mice) and odor sets (see Figure 14 legend for details).

To further investigate the relationship between odor coding and reward coding, we compared our ability to decode reward with our ability to decode the identity of individual odors from the same set of neurons. For this analysis, we used sets of four odors (two rewarded and
Figure 13: Distributions of recorded cells by odor. The number of cells recorded in each area, pPC (a) and OT (b), during sessions in which the odor was familiar (i.e. learned in previous sessions). In total, 42 different odors were used across all mice. The dataset for each mouse consisted of a variety of odors in order to study reward coding beyond the coding for odors alone. Below the histograms of the total number of cells collected in each area are breakdowns of the distributions of cell-odor pairs for each mouse. Specifically, the mice for which many cells were recorded with at least the same 4 odors are shown. These data were used for subsequent analysis of odor-specific decoding.
Figure 13 (Continued)

(a) pPC

(b) OT

Mouse S

Mouse W

Mouse Z

Mouse K

Mouse L

Mouse M

Mouse N

Number of Cells

Cells

Odor

Cells

Odor

Cells

Odor

Cells

Odor

Cells

Odor

Cells

Odor

Cells

Odor

Cells

Odor
Figure 14: Training and testing linear classifiers on different sets of odors. Support vector machines with a linear kernel were trained on the single-trial responses to one set of 4 odors and tested on a different set of 4 odors. Each line corresponds to the test performance for a different mouse or set of test odors. pPC (yellow): 3 curves total; each curve is a different mouse (also, each mouse had a different set of test odors than the other two mice). OT (blue): 5 curves total: 3 curves for three mice with the same test odor set, 1 curve for a second set of test odors for one of the three mice, and 1 curve for a fourth mouse with a different set of odors than the other three mice.

two unrewarded) for which we had the most cells recorded during sessions in which all four odors were used. By choosing the number of odors to be four for this analysis, there were substantially more cells available for each set of odors (mean: 24 cells, range: 12-45 cells) than the previous analysis which required the same eight odors to be used in every session to be included in the dataset. As in all of the previous decoding analyses, reward information was quantified as the decoding accuracy comparing the single-trial responses to rewarded versus unrewarded odors. Additionally, we quantified odor-specific information as the decoding accuracy when classifying responses to one odor versus responses to the other three odors (i.e. “one versus rest” classification). For each area, we tested separately the performance of different mice and sets of odors. In OT, decoding of reward was always as good or better than
Figure 15: Reward and odor coding in OT. Each panel shows decoding performance for four odors from OT cells in a single mouse pooled across sessions in which the odor panel included all four of the odors. Reward coding (blue bars) was quantified as the ability of a linear SVM to classify single-trial responses to two rewarded odors versus two unrewarded odors, in the same way as the previous analyses of population decoding of reward. Additionally, decoding performance is shown for each odor (red bars: rewarded odors, black bars: unrewarded odors). Decoding performance for single odors was determined by training and testing a linear SVM to classify single-trial responses to the odor versus responses to the other three odors (“one versus rest” classification). Chance performance (50%) for all classifiers is indicated with the dashed lines. Error bars are standard error of the mean.
Figure 15 (Continued)
Figure 16: Reward and odor coding in pPC. See description for Figure 15.
decoding of any individual odor (Figure 15). In pPC, however, decoding of reward was substantially lower, at or near chance levels (50%; Figure 16). Interestingly, the one case of strong decoding accuracy for reward in pPC (Mouse Z, Odor Set 1), there was a disparity in coding for individual odors in which the two rewarded odors were coded much more strongly than the two unrewarded odors. This suggests that the reward decoding in this case simply resulted from stimulus coding alone for this set of odors and the particular set of cells that were recorded. Overall, decoding of individual odors was also much lower in pPC than OT, perhaps due to the sparse odor responses in pPC (Figure 8c) resulting in few or no responses to a given odor within the relatively small populations of cells available for this analysis (mean: 32 cells per odor set in pPC). Overall, these analyses demonstrate an explicit code for reward in OT but not pPC.

We then attempted to determine why decoding performance for reward was better in OT than pPC even when the reward labels were shuffled (i.e. Figure 12d,e). First, we considered the possibility that higher sparseness of odor responses in pPC than OT made decoding reward more difficult in pPC than OT with the same number of cells. To minimize the effect of differences in sparseness across the two areas, we repeated the decoding of reward analysis using only cells with a lifetime sparseness between 0.3 and 0.6. We chose this range because represented the greatest overlap between the two areas (Figure 8c). Decoding from OT was still better than pPC, for both un-shuffled (Figure 17a) and shuffled (Figure 17b; p=0.00095 for 2-sample t-test between results for decoding from the entire set of 30 neurons) reward labels, indicating that the better decoding from OT cannot be accounted for by a difference in sparseness alone. We then considered the possibility that the residual difference between OT and pPC decoding performance was due to a difference in the magnitude of the differences in responses across the eight odors. That is, perhaps the responses to each of the eight odors were very different from one another for cells in one area but more similar for cells in the other
Figure 17: Effects of sensory responsiveness on reward decoding. a) Decoding of reward in OT and pPC (same color scheme as Figure 12) for 2, 4, 6, and 8 odors among cells with lifetime sparseness between 0.3 and 0.6. This was the range of lifetime sparseness in which pPC and OT most overlapped (Figure 8c). OT still performed better than pPC even when lifetime sparseness in the two areas was matched in this way. b) Comparison of the reward decoding performance for 8 odors in OT (blue) and pPC (yellow) for the same data as a but with the reward labels shuffled. Higher performance of OT than pPC indicates that sparseness alone cannot account for better reward coding in OT found in previous analyses. c,d) Histograms of the ratio of variance of responses (first 200ms) across odors to variance for the same odor for pPC and OT, respectively. This provided a measure of the magnitude of the difference in responses to different odors. e) Decoding performance on reward label-shuffled data for cells in OT and pPC that were matched for lifetime sparseness (range: 0.3-0.6) and all had a variance ratio in c,d greater than 2.0. OT still performs better than pPC.
area. To quantify the magnitude of differences in odor responses for each cell, we took as a
direct analogy the calculation of the test statistic for ANOVA. Specifically, we calculated ratio of
the variance in responses across odors to the variance in responses within odor (i.e. same
odor). To obtain decode from sets of cells with comparable ratio of variances, we included only
cells for which this value was greater than 2.0 (Figure 17c,d). With this cutoff, the distributions in
Figure 17c versus Figure 17d were indistinguishable (p=0.1094; Kolmogorov-Smirnov test).
Even when limiting the cell populations for reward decoding in OT and pPC to cells with lifetime
sparseness between 0.3 and 0.6 AND ratio of variances greater than two, OT still outperformed
pPC (Figure 17e; p=0.0321 for 2-sample t-test between results for decoding the entire set of 24
neurons). Therefore, the remaining difference in OT and pPC reward coding is due to a factor
which we have not yet determined.

Finally, we examined the relationship between odor coding at the neural level and
behavioral performance with the same odor. If the coding of reward in OT was used to drive
behavior, we would have expected a stronger relationship between coding and behavior (i.e.
“Consistency”) [Majaj 2015] in OT than other candidate areas, such as pPC. To this end, we
plotted the decoding accuracy for individual odors against the behavioral performance of the
mouse with that odor for each area (Figure 18). Indeed, the relationship between odor coding
and behavior was stronger in OT than pPC ($r^2=0.159$ in OT, $r^2=0.040$). However, decoding
performance for most odors in pPC was near chance (50%). This may have precluded us from
uncovering a true relationship between odor coding and behavior in pPC. Sparse odor
responses in pPC might mean that we need to record substantially larger populations of
neurons in order to make an accurate measurement of odor coding, and thus consistency for
pPC.
3.7. **Multiplexing of odor and reward codes**

Together, the previous analyses revealed a multiplexing of odor and reward codes in the olfactory system. To gain further insight into this multiplexing, we established a method to visualize odor and reward selectivity at the same time. For every neuron, we created a bar code that separates responses to each of the eight individual odors but also distinguishes odors based on their reward valence (Figure 19a,b). The position of each odor in the bar code corresponds to the rank-order of the response magnitude (during the first 200ms window), from most positive to most negative. Each position was then color coded according to the valence of the odor, red for rewarded odors and black for unrewarded odors. Finally, the intensity of the color was scaled to normalize the odor response to the strongest response among the eight odors. In this way, a weak odor response is indicated in white whereas a strong odor response is a deep black or red, depending on the valence of the odor.

When the bar codes of all neurons in a brain region are stacked, several features of odor and valence selectivity in the population become apparent. For example, a hypothetical population of neurons in which each is excited selectively by one odor out of eight regardless of
Figure 19: Multiplexing of odor and reward codes. a) Single-cell peri-stimulus time histograms that are used for the barcoding example in b. b) Method to produce barcodes that illustrate both odor and reward sensitivity of a cell’s response during the first 200ms of odor presentation, beginning with the first inhalation. Odors are rank ordered from most positive-going response to most negative-going response, then colored according to valence (red for rewarded, black for unrewarded) and the normalized response to the odor. c,d) Stacks of barcodes for cells recorded in each area during sessions in which all eight odors in the stimulus panel had been successfully learned in a previous session (i.e. familiar odors). e,f,g) The average (mean) bias in rank order for rewarded versus unrewarded odors across cells in each area. The histograms (black bars) show the results for 10,000 datasets generated by shuffling the reward valence labels of each odor response, and the red bar shows the average bias for the actual data in each area. Mean bias statistics for the OT dataset with all cells that perfectly categorize reward valence (i.e. top and bottom rows in d) removed.
valence would appear as a single column of red and black entries on the far left of the stack. In contrast, an area that contains cells that all densely respond to all eight odors and perfectly discriminate valence would appear as a 4-square checkerboard; the checkerboard results because a neuron that perfectly discriminates rewarded versus unrewarded odors could be more excited by all rewarded odors than all unrewarded odors, or vice-versa. The actual data for our pPC and OT recordings are shown in Figure 19c and Figure 19d, respectively.

We then wanted to quantify the significance of the apparent valence selectivity present in the rank ordering of OT odor responses. We calculated the average (mean) valence bias in the center of mass of rewarded versus unrewarded odor ranks. For comparison, we also calculated mean valence bias for datasets with the valence labels shuffled. Consistent with our previous results, the mean valence bias was highly significant for OT (p<0.0001; Figure 19f) but not significant for pPC (p=0.293; Figure 19e). We considered the possibility that valence selectivity in OT arose from a distinct and dedicated population of neurons. To test this possibility, we removed the subset of neurons that were perfectly categorical from the OT data; the mean valence bias was still highly significant among neurons with mixed odor and reward selectivity (p<0.0001; Figure 19g). This visualization and analysis neatly recapitulates the main results of the previous sections. Overall, pPC neurons have sparse selectivity for odors and do not discriminate valence, whereas OT neurons respond much more densely to odors and partially, but not entirely, discriminate valence.

3.8. **Novel odor learning**

The results described thus far established that OT neurons had an explicit representation for reward valence that was multiplexed with odor tuning, whereas pPC neurons lacked an explicit representation for reward valence. Next, we investigated the evolution of
reward coding during the process of learning novel odor-reward associations. Experiments with novel odors allowed us to extend our results, and the results of others [Gadziola 2015; Gadziola 2016], in three important ways. First, we sought to compare the relationship between the achievement of successful behavioral performance with an odor and the emergence of reward coding for that odor in OT. Second, we wanted to obtain error trials to dissociate the relationship of activity in OT to stimulus identity versus motor behavior. Errors were extremely rare for well-learned odors, but necessarily common for novel odors. Finally, we aimed to determine whether the coding for over-trained odor-reward associations differed from that of newly-learned associations.

Approximately every other session following successful learning of the task structure (total of 71 sessions), we removed four of the eight odors that had been used in previous sessions and replaced them with four novel odors. Specifically, two novel rewarded odors replaced two familiar rewarded odors and, likewise, two novel unrewarded odors replaced two familiar unrewarded odors. Mice were not exposed to the odors prior to training and reward valence was randomly assigned to each odor for every mouse. A reward valence prior may have arisen due to an intrinsic association for that odor (e.g. sweet smell) or an association inherited from a similar smelling previously-learned odor. Random assignment of reward valence was designed to minimize the influence of priors on task performance. As described below, mice almost always learned the assigned reward valence, indicating that any existing priors were overridden. In the next section, analysis of neural recordings was restricted to odors with behavioral performance that did not start at a high level in order to avoid confound from odors for which learning was not required. We were able to observe behavior and neural activity throughout the course of learning odor-reward associations since mice typically reached near perfect performance within a single session with a novel odor. The plots in Figure 20a,b show the progression of behavior for each odor during the first session a mouse experienced it; go
Figure 20: Learning of novel odor-reward associations. **a,b)** Go/no-go behavioral responses of mice to novel odors, blue for lick and black for no-lick. Each row is one novel odor. The length of each row corresponds to the total number of presentations of that particular odor during its first session in the odor panel. **c,d,e)** Histograms of the number of licks trials for the first ten, second ten, and third ten presentations of a novel odor illustrate learning of novel odor-reward associations. Red bars indicate responses to rewarded odors and black bars indicate responses to unrewarded odors.
trials are indicated in blue and no-go trials are indicated in black. Correct go responses to rewarded odors become frequent earlier than correct no-go responses to unrewarded odors. For a substantial fraction of rewarded odors, mice licked on ten out of the first ten trials (Figure 20c). A similar frequency of flawless performance on unrewarded odors (i.e. ten out of ten no-go responses) did not occur until twenty trials later (Figure 20d,e). The asymmetry of reward in the go/no-go paradigm might have caused this disparity in the speed of achieving good performance between rewarded and unrewarded odors. A bias toward licking increases the probability of receiving reward under uncertainty, and positive reinforcement encourages the animal to maintain this bias rather than override it (and stop licking). Ultimately, mice never continued to lick for unrewarded odors for multiple sessions. On the other hand, there were some rewarded odors which mice failed to learn. Again, the asymmetry of the go/no-go paradigm may have prevented mice from discovering the true reward association of odors for which their response was no-go. In other words, if the mouse had no reason to suspect that an odor would be rewarded, even if it could easily discriminate that odor from the other odors, the mouse might have assumed that the odor would be unrewarded (for any number of potential reasons) and never tried. As a result, it was not clear whether failure to achieve good performance on these odors reflected 1) an inability to discriminate the rewarded odor from the unrewarded odors or 2) a lack of information regarding the true reward valence label for the odor. Overall, aside from these rare cases, mice demonstrated a flexible ability to learn arbitrary odor-reward associations in a single session.

3.9. Evolution of OT responses during learning of novel odor-reward associations

To determine whether odor-reward associations were reflected in the activity of OT
neurons during learning, we leveraged the fact that only four of the odors were novel in learning sessions and the other four odors were familiar (i.e., learned in a previous session). Responses to familiar rewarded versus unrewarded odors provide some information about the sensitivity of cells recorded during learning sessions to reward. In particular, those responses reveal the steady-state reward selectivity that, presumably, the novel odors will also evoke once learned. Therefore, we wanted to compare the evolution of reward selectivity of novel odors using the familiar odors as a reference point. To this end, we trained a linear classifier (i.e., SVM with a linear kernel) to classify single-trial responses (mean firing rate response in first 200ms of odor sampling) to familiar rewarded versus unrewarded odors. Then, we tested the classifier performance on single-trial responses to novel rewarded and unrewarded odors (Figure 21). Specifically, the data shown in Figure 21 are the average performance of 100 classifiers each trained on a random group of 20 cells out of the 73 recorded during learning in our dataset. We chose to train on 20 cells because classifiers trained and tested entirely on familiar odors perform at approximately 90% correct with this number of cells (Figure 12c). This approach

Figure 21: Decoding reward selectivity from OT responses to novel odors. A classifier was trained to discriminate two familiar rewarded odors from two familiar unrewarded odors, then tested on its ability to decode reward valence from responses to two novel rewarded and two novel unrewarded odors. Error bars show standard error of the mean.
allowed us to generate and test a large collection of classifiers with performance that had not saturated (i.e. 100%). This permitted us to test whether decoding performance, and thus reward selectivity, improved throughout the course of the first session. Indeed, there is a highly significant increase in decoder performance as a function of the presentation number (i.e. trials with the same odor; $p<10^{-6}$), starting near chance performance at the beginning of the session and reaching strong performance (>90%) by the end of the very first session with the novel odors. Overall, these results establish that reward selectivity in the OT develops in parallel with learning at the behavioral level.

We then tested whether detectable changes in the responses of individual OT neurons were present during learning that might explain the increase in reward selectivity that was demonstrated by the decoding analysis. For each OT cell recorded during learning, we calculated the change in response to each novel odor separately, normalized to the mean and standard deviation of the cell's responses to the four familiar odors used in the same session. We then plotted these response changes against the reward valence auROC of the cell to familiar odors (Figure 22). We did not find a significant relationship between reward selectivity

![Figure 22](image) **Figure 22:** Changes in OT cell responses to novel odors during learning. The change in each cell's response to novel odors during the first session of experience is plotted against the reward valence auROC of the cell for familiar rewarded versus unrewarded odors. The data are split to show response changes to novel rewarded (a) and novel unrewarded odors (b).
3.10. **Analysis of error trials**

The go/no-go task structure tied the reward valence with the motor decision. That is, rewarded odor trials involved a go response and unrewarded odor trials involved a no-go response, given that the animal performed correctly. In order to dissociate effects of the reward valence from the motor decision in neural activity, we examined OT responses on error trials (Figure 23). Due to the asymmetry in hit rate versus correct rejection rate in our task, there were too few misses to be used for this analysis; instead, we focused on false alarm trials. The effect of the reward valence was quantified by calculating the auROC of OT neuron responses across hit versus false alarm trials (“valence auROC”), for which the motor decision was the same but the valence differed. Likewise, the effect of the motor decision was quantified by calculating the auROC of OT responses across correct rejection versus false alarm trials (“motor auROC”), for which the valence was the same but the motor decision differed. Importantly, we never included responses from the same odor among both the false alarm trials and correct rejection trials when calculating the motor auROC. This ensured that odor tuning did not artificially decrease the motor auROC relative to the valence auROC since the odors were always different (by definition) between hit and false alarm trials. We did not find a significant difference in the
pairwise comparison between valence and motor auROCs (p=0.344). We concluded from this result that the reward valence tuning we have observed did not strictly reflect a difference in motor behavior.

3.11. Summary of results

Mice learned to successfully perform an odor-reward categorization task over the course of a few days. Once the task structure had been learned, mice could flexibly apply the task rules in order to acquire novel odor-reward associations within a single session. This categorization task allowed us to study how odor and reward coding interact in the olfactory system. We observed a transformation from a code in pPC that had sparse odor responses and weak reward selectivity to a code in OT that had strong reward selectivity and less sparse odor responses than pPC. Reward selectivity in OT emerged in the pre-decision time during the first
sniff of odor cues and was detectable, though weak, at the single-cell level. Moreover, reward selectivity and successful behavioral performance developed in parallel as novel odor-reward associations were learned. These results provide important constraints on how odors are processed in the olfactory system during reward-motivated behaviors. Several implications of these results will be discussed in the next chapter.
4. **Conclusion and Discussion**

The posterior piriform cortex (pPC) and olfactory tubercle (OT) are two high-level olfactory areas situated at the intersection of sensory and reward-related brain regions. Unfortunately, little is known about odor coding, or how reward information contributes to odor representations, in either of these areas. Here, I have shown that there is a transformation from a code in pPC that has strong odor selectivity but weak reward selectivity to a code in OT that has strong reward selectivity. Consistent with strong odor selectivity, responses in pPC are substantially more modulated by sniffing than responses in OT. In the OT, odor representations are multiplexed with reward information in the activity of single neurons. Reward selectivity is present in the OT during the pre-decision time that starts with the first sniff of odor cues. This reward selectivity emerges in parallel with successful behavioral performance as novel odor-reward associations are learned. The following sections discuss some of the potential interpretations and implications of these results.

### 4.1. How do OT neurons acquire reward valence information?

OT is heavily interconnected with the reward system. In fact, reciprocal connectivity between the OT and ventral tegmental area (VTA) dopaminergic neurons is the strongest in the mammalian olfactory system: OT is the strongest source of input among olfactory brain regions and is the major olfactory recipient of dopaminergic input [Wesson 2011; Watabe-Uchida 2012]. The presence of this loop raises the question, does reward-selective activity start in the OT or VTA? Dopamine-induced plasticity during learning could train OT cells to decode reward valence from feedforward sensory input. However, the prevalence of reward-selective activity in
our OT data, 46% of cells during the pre-decision period, seems more plausibly explained by broad input from VTA dopaminergic neurons (or another source) that is already reward-selective than each cell in OT decoding valence separately. An intriguing possibility is that a small population of OT neurons that are appropriately tuned for the odor decode reward valence, pass this information along to VTA, and then a non-odor-selective signal from VTA biases the activity of the entire OT population (among other areas). Since the VTA is only one synapse from OT, and feedback from VTA only two synapses away, the delay between decoding in OT and the onset of ubiquitous reward activity would be very short. As a result, determining which comes first is difficult to address with recordings alone.

An important experiment for future work would be to “break the loop” between OT and VTA. Inactivating dopaminergic input to the OT could reveal how much reward-selectivity is generated locally, or at least independent of VTA, and how much is inherited. Ideally, this could be done temporarily during the beginning of odor sampling and only on a small fraction (say, 10%) of trials. Inactivating the axons of VTA inputs to prevent dopamine release locally without disrupting VTA activity generally would be preferable. However, current inhibitory optogenetic tools, such as archaerhodopsin or halorhodopsin, might not be capable of sufficiently shunting action potentials in axons and axon terminals. Therefore, a failure to eliminate reward activity during the beginning of odor responses in OT would yield an ambiguous result. On the other hand, a positive result would be a clear confirmation that a substantial portion of the reward selectivity is inherited from VTA. Conversely, inactivation of the OT paired with recordings of dopaminergic neurons during behavior could be used to assess the contribution of OT input to olfactory-cued reward responses in the VTA.

Of course, VTA is not the only potential source of reward information to the OT. Other neuromodulatory inputs to OT, such as serotonin, could carry this information as well. Most notably, the amygdala and orbitofrontal cortex are potential candidates. These areas have
access to sensory information through inputs from multiple olfactory regions and have substantial reward-related activity during tasks similar to ours [Schoenbaum 1998; Schoenbaum 1999; Ramus 2000; Roesch 2007]. Importantly, neurons in orbitofrontal cortex and basolateral amygdala also acquire selectivity for odor-reward associations early during learning [Schoenbaum 1998]. However, basolateral amygdala primarily responded to negative valence associations, making orbitofrontal cortex the more likely candidate to play a role in the task presented here which involved only positive and neutral valences. Similar to the case of dopaminergic input, these areas likely contribute to the reward activity present in OT and assessing the strength of those contributions will be important, but challenging.

4.2. How do OT neurons acquire reward-selective responses so early during learning?

The emergence of reward-selective responses in a part of the striatum, such as the OT, is not unprecedented. Monkeys trained to associate reward with visual stimuli develop reward selectivity in the dorsal striatum that is apparent in the activity of individual neurons within a few trials of successful learning [Schultz 2003; Pasupathy 2005]. In contrast, neurons in the prefrontal cortex took much longer to acquire reward selectivity [Pasupathy 2005]. In this context, the results presented here strengthen the analogy between OT and dorsal striatum based on anatomy, cell types, and cytoarchitecture that was discussed in the introductory chapter. OT might simply be the part of striatum that is largely devoted to processing olfactory information but otherwise implements computations that are similar to the rest of striatum [Wesson 2011; Giessel and Datta 2014].

The provocative interpretation is that reward-selective responses emerge in striatum early during learning because plasticity in the striatum supports decision-making. Indeed,
corticostriatal inputs have been shown to drive decision-making in an auditory discrimination task [Zmenskiy]. A followup study by the same group showed that plasticity occurred at corticostriatal synapses over the course of training [Xiong 2015]. Potentially, the role of striatum in sensorimotor decision-making extends to the OT.

4.3. **Is reward activity definitely absent in olfactory cortex (e.g. pPC) during decision-making?**

The analyses presented here only looked for an explicit code for reward valence. Partly, this strategy was taken due to experimental limitations. Our tetrode recordings only yielded one or a few cells per session (and sometimes none), which makes it difficult to detect more subtle implicit codes. Furthermore, the high sparseness of odor responses in pPC might mean that much larger populations of neurons than the dataset we collected will be necessary to reveal more complicated structure. Alternatively, the short time required for decision-making in this task (200-300ms) might limit the usefulness of highly intricate coding schemes, or at least ensure that the coding scheme involves differences in firing rate between rewarded and unrewarded odors that are detectable with our analyses. However, I do plan to analyze higher order correlations in the future to see if these limitations are surmountable, and hope that future studies utilizing high-density recording techniques will shed light on this issue.

4.4. **What gives rise to the disparity in lifetime sparseness between pPC and OT?**

Piriform cortex, including pPC, receives olfactory bulb input primarily from mitral cells whereas olfactory tubercle receives input primarily from tufted cells [Neville and Haberly 2004]. Tufted cells are more broadly tuned than mitral cells [Mori and Sakano 2011] and, in turn, pPC
and OT might inherit this tuning disparity from their inputs. Sparseness also depends on the relative convergence of sensory inputs to each area, which is not known with sufficient precision to determine its contribution. Finally, local processing can also play a role. The pPC has extensive recurrent connections which shape odor tuning, whereas lateral connectivity in the OT is substantially lower. The factors that contribute to sparseness in these areas, and the effects this disparity has on sensory processing, should be studied further.

4.5. How to dissociate the contributions of reward valence versus motor decision in neural activity?

The go/no-go paradigm links reward valence and motor decision, making it difficult to determine the relative contributions of each to neural activity. In contrast, the two-alternative forced choice (2-AFC) paradigm dissociates the reward valence from the motor decision while still requiring only a binary choice. Rodents have been successfully trained on odor discrimination tasks with 2-AFC [Uchida and Mainen 2003]. Furthermore, 2-AFC does not have a reward asymmetry like go/no since both choices can be rewarded equally. Symmetrically rewarded responses make it easier to interpret behavior near perceptual threshold, which is required to generate sufficient error trials to measure choice probability [Britten 1996]. This approach could help us better understand the role of the OT in decision-making.

4.6. Final thoughts

The work presented here sheds light on the representation of a behaviorally relevant aspect of the environment, reward, and how it relates to sensory coding in the olfactory system. However, many important issues remain to be addressed in future work, such as whether and how reward coding in sensory cortex contributes to learning and decision-making, the potential
for more subtly-structured implicit coding schemes that those examined here, and how activity in these areas interacts with other brain regions to drive behavior. New tools that permit recording large ensembles of neurons simultaneously or inactivating specific populations of neurons could help tremendously in this effort. Genetic tools available in mice will hopefully allow the microcircuitry that underlies the phenomena described here to be dissected with cellular-level precision. Studies of neural coding during olfactory behavior with a high degree of experimental control are not new, but certainly still at an early stage. I am very excited to see how the field develops.
Bibliography


Davison IG and Katz LC (2007) Sparse and Selective Odor Coding by Mitral/Tufted Neurons in


Ikemoto S, Witkin BM and Morales M (2003) Rewarding injections of the cholinergic agonist carbachol into the ventral tegmental area induce locomotion and c-Fos expression in the
retrosplenial area and supramammillary nucleus. *Brain Res.* 969: 78-87.


Rate Coding and Decorrelated Population Activity. *Neuron* 74: 1087-1098.


Pratt WE and Mizumori SJY Characteristics of basolateral amygdala neuronal firing on a spatial memory task involving differential reward. *Behav. Neurosci.* 112: 554-570.


Roman FS, Simonetto I and Soumireu-Mourat B (1993) Learning and memory of odor-reward


Slotnick B and Berman EJ (1980) Transection of the lateral olfactory tract does not produce


Suzuki NS and Bekkers JM (2006) Neural coding by two classes of principal cells in the mouse


Wesson DW and Wilson DA (2011) Sniffing out the contributions of the olfactory tubercle to the


