Report

Neuron

Odor Perception on the Two Sides of the Brain: Consistency Despite Randomness

Highlights

- A random model predicts observed preservation of correlations in piriform responses
- The model supports consistent agreement about odor quality among individuals
- Consistent generalization may require the full complement of piriform neurons

Authors

Evan S. Schaffer, Dan D. Stettler, Daniel Kato, Gloria B. Choi, Richard Axel, L.F. Abbott

Correspondence

ess2129@columbia.edu

In Brief

Random connectivity from the olfactory bulb to piriform implies that the representation of odors will differ in different cortices. Schaffer et al. demonstrate in a model that consistent agreement about odor quality occurs but requires the full complement of piriform neurons.

Schaffer et al., 2018, Neuron *98*, 736–742
May 16, 2018
https://doi.org/10.1016/j.neuron.2018.04.004

CellPress

CellPress

Neuron
Report

Odor Perception on the Two Sides of the Brain: Consistency Despite Randomness

Evan S. Schaffer,[1,7,*] Dan D. Stettler,[1,5] Daniel Kato,[1] Gloria B. Choi,[1,6] Richard Axel,[1,2,3] and L.F. Abbott[1,4]
[1]Mortimer B. Zuckerman Mind Brain Behavior Institute, Department of Neuroscience, Columbia University, New York, NY 10027, USA
[2]Department of Biochemistry and Molecular Biophysics, Columbia University, New York, NY 10032, USA
[3]Howard Hughes Medical Institute, Columbia University, New York, NY 10032, USA
[4]Department of Physiology and Cellular Biophysics, Columbia University, New York, NY 10032, USA
[5]Present address: Kallyope, Inc., 430 East 29th Street, 10th floor, New York, NY 10016, USA
[6]Present address: McGovern Institute for Brain Research, Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, Cambridge, MA 02139, USA
[7]Lead Contact
*Correspondence: ess2129@columbia.edu
https://doi.org/10.1016/j.neuron.2018.04.004

SUMMARY

Neurons in piriform cortex receive input from a random collection of glomeruli, resulting in odor representations that lack the stereotypic organization of the olfactory bulb. We have performed *in vivo* optical imaging and mathematical modeling to demonstrate that correlations are retained in the transformation from bulb to piriform cortex, a feature essential for generalization across odors. Random connectivity also implies that the piriform representation of a given odor will differ among different individuals and across brain hemispheres in a single individual. We show that these different representations can nevertheless support consistent agreement about odor quality across a range of odors. Our model also demonstrates that, whereas odor discrimination and categorization require far fewer neurons than reside in piriform cortex, consistent generalization may require the full complement of piriform neurons.

INTRODUCTION

Olfactory perception involves both odor discrimination and generalization. Discrimination relies on the detection of differences between odors, whereas generalization requires the identification of similarities and dissimilarities across odors. The ability to generalize is vital to performance in any olfactory task because natural variability implies that the same stimulus is never experienced twice. Perceptual consistency across odors requires that generalization be similar among individuals. We thus explore whether the unstructured representation of odors in piriform cortex can support consistent generalization.

Individual olfactory sensory neurons in the mouse express one of 1,500 receptor genes (Buck and Axel, 1991; Godfrey et al., 2004; Zhang and Firestein, 2002). Neurons expressing a given receptor project with precision to two spatially invariant glomeruli within the olfactory bulb (Mombaerts et al., 1996; Ressler et al., 1993, 1994; Vassar et al., 1993, 1994). Odors elicit distributed neural activity in the sensory epithelium that is transformed in the olfactory bulb, where each odor evokes a distinct spatial pattern of glomerular activity (Davison and Katz, 2007; Igarashi and Mori, 2005; Ma et al., 2012; Meister and Bonhoeffer, 2001; Soucy et al., 2009; Takahashi et al., 2004). A second transformation occurs in the piriform cortex, where individual odors activate unique ensembles of neurons that lack discernable spatial patterning (Illig and Haberly, 2003; Iurilli and Datta, 2017; Poo and Isaacson, 2009; Rennaker et al., 2007; Stettler and Axel, 2009; Sugai et al., 2005; Zhan and Luo, 2010). Representations of individual odors are distributed across the entire piriform with no apparent large-scale (Stettler and Axel, 2009) or local spatial structure (Figure S1). One model consistent with both the anatomy and the physiology assumes that each piriform neuron samples a random combination of glomerular inputs (Choi et al., 2011; Davison and Ehlers, 2011; Ghosh et al., 2011; Miyamichi et al., 2011; Sosulski et al., 2011). Random connectivity implies that the piriform representation of a given odor will differ among different individuals and across brain hemispheres in a single individual. Does random connectivity from the olfactory bulb to piriform cortex support generalization and allow shared experiences to "align" the perception of odor quality in different individuals across a wide range of odors?

We have combined mathematical modeling with analysis of *in vivo* optical imaging of odor responses in piriform cortex and find that many properties of the piriform odor representation are in good accord with predictions of a model with random input. Consistent with previous models (Babadi and Sompolinsky, 2014; Barak et al., 2013; Cho and Saul, 2009; Litwin-Kumar et al., 2017), we observe that piriform representations simulated with random input are less correlated than the bulb representations but maintain sufficient correlation between odors to support generalization. We consider how individuals with different piriform odor representations can agree consistently about odor quality across a range of odors and how conflicts are avoided between the two sides of a single individual's brain. Our model demonstrates that whereas odor discrimination and categorization require far fewer neurons than reside in piriform cortex (Babadi and Sompolinsky, 2014), consistent generalization requires the full complement of piriform neurons.

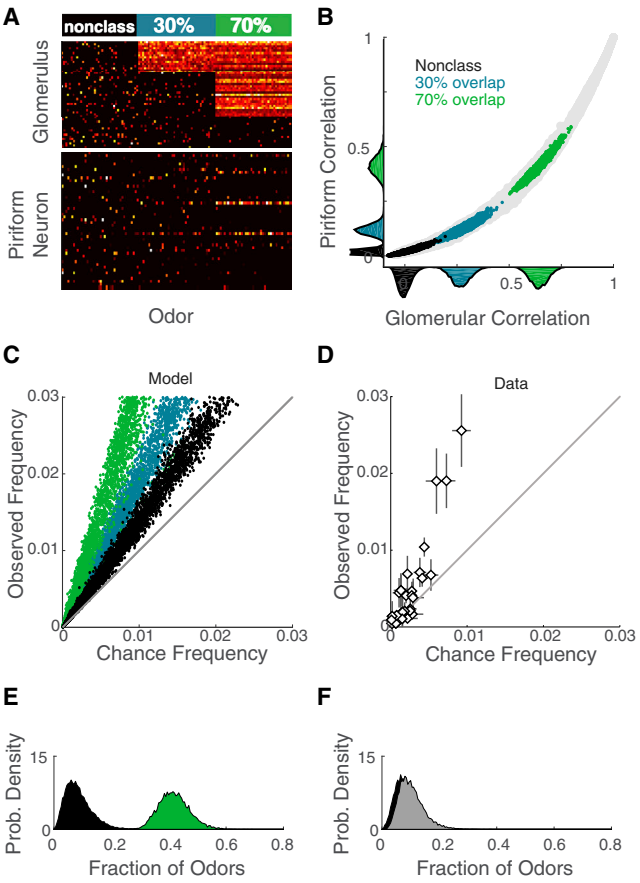

Figure 1. The Transformation from Model Bulb to Model Piriform Partially Preserves Stimulus Similarity

(A) The representation of a model odor in the bulb is transformed through random connections to a representation in the piriform. Example model odors in the bulb (top) and piriform (bottom), with three different values of the excess overlap parameter f: nonclass odors ($f = 0$, black), slightly overlapping odors ($f = 0.3$, blue), and very overlapping odors ($f = 0.7$, green). For visual clarity, just 50 glomeruli and 50 piriform neurons are shown. Color scale is normalized; white, max response; red, weak response; black, no response.

(B) Correlation in piriform representation versus correlation in glomerulus representation. Each dot corresponds to the correlation coefficient between two odors with the same level of overlap ($f = 0.3$ or $f = 0.7$), color-coded according to the conventions in (A). For reference, the full spectrum of values of f ranging from 0 to 1 is shown in gray. Distributions of correlation in the glomeruli and in piriform are shown along their respective axes.

(C) Observed probability of a piriform cell responding to two odors in the model versus the expected probability if odor representations were statistically independent ($f = 0.0$, $f = 0.3$, and $f = 0.7$ shown in black, blue, and green, respectively). In order to recapitulate the variability in representation sparseness observed *in vivo*, model odors were generated with a range of sparseness values from $S_x = 0.05$ to $S_x = 0.2$.

(D) Same as (C), but for *in vivo* odor responses (average ± SEM).

(E) Class selectivity of a neuron is quantified as the difference in that neuron's mean response to class odors versus nonclass odors. Shown is the probability distribution over the fraction of class (green) and nonclass (black) odors eliciting a response for model neurons in the 4th percentile for class selectivity.

(F) Same as (E), but for a class defined by an arbitrary collection of odors not necessarily activating similar glomeruli. Selected class members in gray and nonclass members in black.

RESULTS

Piriform Responses to Odor Pairs Exhibit Correlations Predicted by a Randomly Connected Model

Similar odors evoke correlated patterns of activity in the olfactory bulb (Davison and Katz, 2007; Igarashi and Mori, 2005; Ma et al., 2012; Meister and Bonhoeffer, 2001; Soucy et al., 2009; Takahashi et al., 2004). Random wiring from the bulb to piriform might suggest that these correlations would be absent in cortex. However, a number of theoretical studies have shown that correlations in a set of inputs are reduced, but not eliminated, when activity is transmitted through random synaptic connections (Babadi and Sompolinsky, 2014; Barak et al., 2013; Cho and Saul, 2009). Indeed, piriform responses to pairs of odors activating similar glomeruli can show considerable overlap (Figure 1D; octanal and hexanal, 24% overlap in piriform [Stettler and Axel, 2009]; 60%–70% overlap in bulb [Meister and Bonhoeffer, 2001; Igarashi and Mori 2005]; see also 26% in bulb [Ma et al., 2012]).

To explore the implications of random connectivity, we constructed a model in which the connectivity between the olfactory bulb and piriform cortex is random (STAR Methods). The model has an input layer of $N_x = 1,000$ glomeruli and a cortical layer with N_y neurons. Model odors activate a sparse ensemble of S_xN_x glomeruli, with sparseness $S_x = 0.1$, unless otherwise noted. Bulb activity both excites and inhibits the cortical layer in a balanced manner, so that the average bulb-derived input across the cortex is zero. Connectivity from bulb to piriform is sparse ($S_c = 0.2$), so each piriform neuron receives input from $S_cN_x = 200$ random glomeruli, and approximately 20 of these are activated by a given odor. Model piriform neuron responses are a threshold-linear function of the sum of their inputs with the threshold chosen so that, on average, S_yN_y cortical neurons respond to an odor with $S_y = 0.06$, matching the sparseness of responses observed *in vivo*. Firing rates of the active glomeruli are chosen from a lognormal distribution. We defined classes of model odors with a mean shared fraction f of active glomeruli (STAR Methods), and we focus our analysis on three sets of odors (Figure 1A), defined by $f = 0$ (nonclass), $f = 0.3$ (weak class), and $f = 0.7$ (strong class).

Correlations between simulated piriform activities for odor pairs are related to but smaller than correlations of glomerular activities (Figure 1B), in agreement with previous results (Babadi and Sompolinsky, 2014; Cho and Saul, 2009). The shape of this curve follows from the observation that sparse activity patterns can more faithfully preserve positive correlations than negative correlations. Whereas input correlations close to 1 tend to be largely preserved, negative input correlations are transformed into piriform correlations near zero because uncorrelated sparse representations have very little overlap, leaving minimal dynamic range for further anticorrelation. Consequently, the curve relating input correlation to output correlation has a shallow slope for low input correlation.

Comparison of model results with data is facilitated by considering the fraction of neurons responding to pairs of odors, rather than the correlation. The fraction of neurons responding to pairs of simulated odors, drawn from all three classes, exceeds the value expected from an uncorrelated odor representation

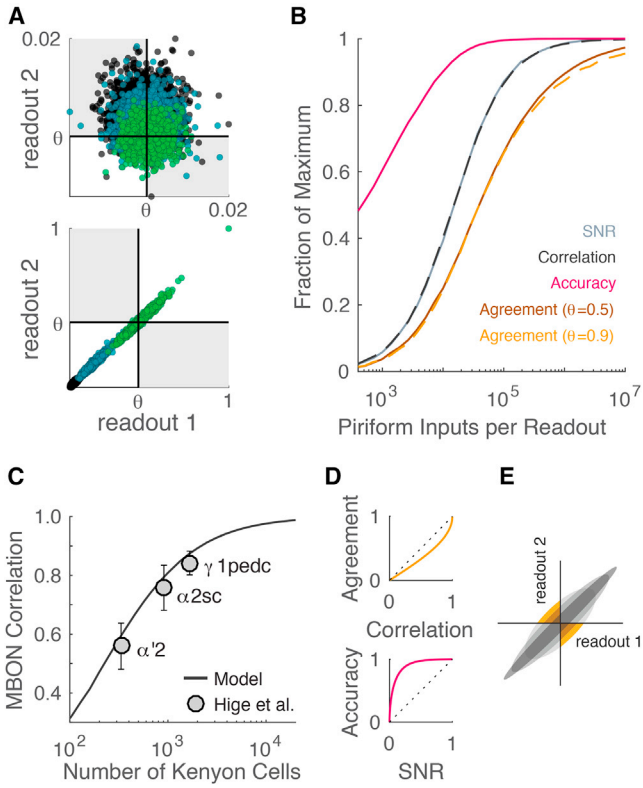

Figure 2. Agreement between Readout Units Requires a Large Piriform

(A) Response of a readout unit versus the response of a second readout connected to an independent set of piriform neurons, either before training (top) or after training to a single odor (bottom), for a panel of odors with 70%, with 30%, or at chance levels of overlap with the trained odor, color coded as in Figure 1. Gray boxes denote the regions in which the readouts do not agree for $\theta = 0.5$ and $f = 0.7$.

(B) Scaling with N_y of readout agreement (with $f = 0$) versus other measures of readout performance: readout agreement with a threshold of $\theta = 0.5$ (brown) or $\theta = 0.9$ (orange), readout correlation (black), SNR (gray), and accuracy (magenta). All quantities except SNR are defined from 0 to 1 (Max[SNR] = 3.4). We normalize SNR such that the minimum value is 0 and the maximum value is 1 to enable comparison to other quantities.

(C) Comparison with experimental data from the *Drosophila* mushroom body. The correlation between two model MBONs as a function of the number of inputs they receive (black). Published data (Hige et al., 2015b) showing the correlation across odors in three conjugate pairs of MBONs versus the number of KCs that innervate each MBON (gray; mean ± SEM).

(D) Readout agreement is smaller than readout correlation (top), while readout accuracy is greater than normalized readout SNR (bottom).

(E) Intuition for why readout agreement is smaller than readout correlation. As the correlation between two readouts grows, their joint probability distribution becomes elongated (gray), but even when this distribution is very elongated, an appreciable portion remains in the off-diagonal quadrants (orange).

(Figure 1C). For independent representations, the fraction of cells responding to both odor A and odor B would be $p_A p_B$, where p_A and p_B are the probabilities for responses to A or B alone. For simulated odors that activate shared sets of glomeruli (Figure 1C), this is a direct result of input-derived correlations maintained despite random connectivity (Babadi and Sompolinsky, 2014; Cho and Saul, 2009). In addition, more neurons than expected for independent representations respond to odor pairs that do not share active glomeruli (Figure 1C). This correlation arises from common glomerular inputs that occur by chance despite random connectivity (see also Litwin-Kumar et al., 2017).

The model results make two predictions about correlations in the piriform representations of odor pairs. First, the average correlation between molecularly unrelated odor pairs should be positive, not zero. Second, odors that activate similar sets of glomeruli should evoke correlated representations in piriform, but these correlations should be smaller than in the bulb. Two-photon calcium imaging of layer 2 of piriform cortex (STAR Methods) reveals that the fractions of neurons responding to odor pairs (Figure 1D) are consistent with the model (Figure 1C). Most pairs produce larger fractional responses than would be expected for statistically independent activity ($p < 0.02$, Wilcoxon signed-rank test). In our model, correlations in the bulb are the major source of correlations in piriform. Adequate data from imaging of the response to the same odor pairs in both bulb and piriform are unavailable. Therefore, we cannot at present experimentally compare correlations across these two olfactory areas.

Random Connectivity Can Produce Odor-Class Selective Piriform Neurons

We next asked whether piriform neurons that respond selectively to a set of similar odors can arise from random input wiring. We define class selectivity as the difference between the mean responses of a neuron to class and nonclass odors and then ranked the neurons in our model according to this selectivity. We illustrate the results by showing the probability distribution over the fraction of odors eliciting a response for model neurons in the 4th percentile for class selectivity (Figures 1E and 1F). On average, these neurons respond near chance levels ($S_y = 6\%$) to nonclass odors, but they respond to 40% of class odors (Figure 1E). We show that this selectivity reflects overlap in the bulb by constructing a nonsensical class defined as a random collection of 1,000 odors with $f = 0$. Neurons in the 4th percentile of selectivity to this artificial class respond near chance levels to both class and nonclass odors (Figure 1F). Thus, even a randomly connected piriform will have some cells with selectivity to a class of similar odors.

Learning Single-Odor Associations Can Align Population Readouts from Different Piriform Representations

Given the randomness of the input to piriform, the same odor will be represented differently in different brains. We now consider how these individuals can nevertheless align the qualities they assign across a wide range of odors. A similar problem arises in considering the alignment of odor qualities inferred from the two sides of a single individual's brain. To define an odor quality in our model, we introduce a linear readout that characterizes the output of an entire population of piriform neurons. The readout is a weighted sum of piriform activities $z = \sum_i w_i y_i$, where y_i is the activity of piriform neuron i and w_i is its weight. Initially, the readout weights are chosen randomly, and we examine the resulting readout responses from two different randomly wired model piriform cortices (Figure 2A, top). Readout responses obtained from two different model piriform cortices using the same

weights are uncorrelated (correlation coefficient = 0.03, −0.03, and 0.005, for odors with f = 0.7, 0.3, and 0, respectively) because the two piriform representations are completely different.

We next asked whether shared training to a single odor in two different individuals results in readouts that agree across a wide range of odors. In this simulation, we set the weights for each readout equal to the activity produced by the odor being learned in the corresponding piriform ($W_i = y_i^*$, where the piriform response to the trained odor is y_i^*; STAR Methods). This implements a form of Hebbian learning (Oja, 1982). This single shared experience, corresponding to associative learning with a common odor, does an excellent job of aligning the two readouts in response to other odors (Figure 2A, bottom). Simulated odors sharing either 30% or 70% of activated glomeruli with the trained odor evoke significantly larger responses in the readouts than all nonclass odors. Two readouts not only distinguish class from nonclass, but can also rank the similarity of a panel of odors extremely well, even if the odors are not similar to one another or to the trained odor (the $f = 0$ case). Moreover, the responses of the two readouts across all simulated odors are well correlated (correlation coefficient = 0.98, 0.97, and 0.86, for f = 0.7, 0.3, and 0, respectively). It is worth emphasizing that the observation that this works for all odors implies that the identity of the training odor is inconsequential. Thus, after training with a single exemplar odor, two readouts can generalize in a similar way across other odors.

The degree of correlation between two readouts due to single-odor learning depends on the number of piriform neurons. This is of interest because of the observation that the number of piriform neurons exceeds the number required for efficient categorization (Babadi and Sompolinsky, 2014). The readout correlation increases and saturates as a function of the number of piriform neurons (Figures 2B and S2). The correlation coefficient between readouts modified by a shared training experience provides a measure of consistency. The ability of a single readout to categorize odors, on the other hand, is characterized by its signal-to-noise ratio (SNR; Babadi and Sompolinsky, 2014; Litwin-Kumar et al., 2017). As a function of the number of piriform neurons, the SNR saturates (Babadi and Sompolinsky, 2014; Litwin-Kumar et al., 2017), and in our model, this saturation occurs when roughly 100,000 piriform neurons provide input to the readout. It has been noted that this is an order of magnitude less than the actual number of piriform neurons (Babadi and Sompolinsky, 2014). The curve showing the correlation coefficient for two readouts (black curve in Figure 2B) overlaps with the curve for the SNR scaled so that its maximum value is 1 (gray curve in Figure 2B). We discuss the reasons for this overlap in the STAR Methods.

To provide a test of our calculation of readout alignment, we exploit experimental data from the *Drosophila* mushroom body. The fly affords an opportunity to examine genetically defined neurons that readout from the two different mushroom bodies in the two hemispheres of the fly brain. In *Drosophila*, projection neurons connect glomeruli in the antennal lobe (the insect analog of the olfactory bulb) to Kenyon cells (KCs) that form the mushroom body (the piriform analog) (Marin et al., 2002; Wong et al., 2002). As in the piriform, glomerular inputs to the KCs show no apparent structure (Caron et al., 2013; Gruntman and Turner, 2013; Murthy et al., 2008). Further support for the random nature of these connections is provided by electron microscopy (EM) data from a *Drosophila* larva (Eichler et al., 2017), indicating that input connections to the KCs are not only unstructured, but are also completely different on the two sides of the brain. KCs synapse onto mushroom body output neurons (MBONs) that are analogous to the model readout we have introduced. Learned associations are established by modifying the strength of the KC to MBON synapses (Aso and Rubin, 2016; Hige et al., 2015a; Séjourné et al., 2011). Thus, homologous MBONs on the two sides of the brain of a single fly should exhibit the response alignment that we have discussed. Correlations were determined for three MBON types that receive different numbers of KC inputs (Hige et al., 2015b). This allows us to test not only the values of the correlations we computed, but also their dependence on the number of neurons driving the readout.

We constructed a fly analog of our piriform model, adjusting the numbers of neurons and synapses appropriately (STAR Methods). The correlation coefficient between two model MBONs as a function of the number of inputs (Figure 2C) resembles the piriform correlation curve in Figure 2B. The fly model correlation saturates at a value of roughly 2,000 KC inputs, the actual number of KCs in the mushroom body. Hige et al. (2015b) have analyzed the responses of left-right pairs of MBONs that receive different numbers of KC inputs to a panel of odorants. The correlation coefficient between MBON pairs determined experimentally shows a striking match with the predictions of the model (Figure 2C).

Readout of a Sufficiently Large Piriform Population Supports Consistent Choice

Odor-guided behavior typically involves making a binary choice; for example, to act or not to act. We have described a readout that provides a continuous measure of the piriform response. We now model choices by comparing this readout value to a threshold. Readout values greater than the threshold are interpreted as a choice to act; those below threshold as a choice not to act. This allows us to compute the accuracy with which a trained readout can guide choice (Figure 2B; accuracy is defined as fraction correct). Accuracy is an increasing function of the SNR that saturates before the SNR reaches its maximum (Figure 2D, top). As a result, accuracy saturates at a lower number of piriform neurons than does the SNR (Figure 2B), further emphasizing the disparity between the actual number of piriform neurons and the number needed to support categorization and choice.

The threshold we have introduced allows us to investigate the consistency of choice across individuals or brain hemispheres. We trained two readouts, each connected to an independent model piriform cortex (10^6 neurons), on a single odor. We then compared the readouts to a threshold and examined the choice agreement across a large number of nonclass odors. Good agreement between readouts is substantially more difficult to obtain than is accuracy in a typical categorization task such as distinguishing class from nonclass odors. We characterize the value of the threshold with a parameter θ that is equal to the fraction of odors that produce a subthreshold readout response. At a

threshold for which odors are equally divided between supra- and subthreshold ($\theta = 0.5$), 95% of the choices are the same for two trained readouts, but agreement is at chance levels (~50%) for two untrained readouts (Figure 2A). Note that this consistency is observed for odors that are both similar to and different from the trained odor. Thus, readouts receiving input from different piriform representations can support consistent choices after a single shared training experience.

We next asked how the consistency of choices depends on the number of neurons in piriform. For this purpose, we introduce a quantity A_θ called the agreement. In defining the agreement, we correct for the bias introduced by thresholds different from 0.5 by subtracting the fraction of choices due to chance from the fraction of choices that agree. Normalizing the resulting quantity to a maximum value of 1 yields our measure of agreement A_θ (STAR Methods).

Whereas accuracy is determined by SNR, agreement is a function of the readout correlation coefficient and is always the smaller quantity (Figure 2D, bottom). A_θ saturates at higher numbers of piriform neurons than the accuracy, SNR, or readout correlation coefficient (Figures 2B). This is illustrated graphically in Figure 2E. As the number of piriform neurons grows, the values of the two readouts become more correlated and fall in an increasingly elongated elliptical distribution (Figure 2E, gray). For $\theta = 0.5$, readout agreement corresponds to the fraction of the distribution not in the off-diagonal quadrants (Figure 2E, orange). Even when the distribution is very elongated (high correlation), an appreciable portion remains in these quadrants. The slow increase of A_θ as a function of the number of piriform neurons is a general property that is relatively insensitive to the threshold used because A_θ has been corrected for bias and normalized.

Our conclusions concerning the agreement between two readouts hold when we extend our analysis to multiple readouts. We introduce an additional parameter to specify the fraction of similarly responding readouts required to say that the ensemble "agrees" (STAR Methods). For small values of this parameter, the many readout case looks very similar to the two-readout case; for large values, maximal agreement requires at least an order of magnitude more inputs from piriform (Figure S2). Furthermore, the number of piriform inputs required for maximal readout agreement reduces only slightly for odors that more densely activate the olfactory bulb (Figure S2). Thus, a single piriform cortex can support accurate categorization and choice with fewer than 10^5 neurons, an order of magnitude less than the actual number. However, 10^6 piriforms are required for choice agreement between different individuals or across the two sides of the brain. This may provide a rationale for the large number of piriform neurons.

DISCUSSION

Neurons in piriform cortex receive input from a random collection of glomeruli, resulting in an odor representation that lacks the stereotypy of the olfactory bulb. Random connectivity implies that each piriform cortex is unique, posing the problem as to how different individuals generalize consistently across stimuli. Generalization implies that different individuals, after learning a conditioned response to odor A, will generate the learned behavior in response to odors with glomerular representations similar to A, but not to odors that activate different glomeruli. Consistent generalization is also necessary across the two differently wired piriform cortices of a single individual if inter-hemispheric conflicts are to be avoided. Our model demonstrates that readouts from two different randomly wired cortices can be highly correlated after a single shared experience. In this manner, perceptual consistency will not require individuals to have lived identical lives. The degree of correlation between model readouts depends on the number of randomly wired neurons, generating predictions that are in excellent agreement with data from the mushroom body output neurons of the fly. Moreover, the ability of the model piriform to support consistent choices across individuals requires that the piriform cortex contain a large number of neurons, a number in accord with that observed in the mouse. Modeling studies using more traditional performance measures such as dimensionality and readout SNR have suggested that piriform has an excess of neurons (Babadi and Sompolinsky, 2014). We find that consistent decision-making may require the full complement of piriform neurons.

The model we have considered is based on an assumption of random wiring. We investigated this assumption further by examining *in vivo* data for evidence of structure. We could detect no local spatial structure, no periodic structure, and no clustering across the piriform surface. We did find positive correlations in the responses of individual neurons to odor pairs, but these are consistent with expectations given random wiring from the bulb to piriform. The fact that piriform odor representations generated by random connectivity can "inherit" correlations from bulb representations is essential for generalization across odors.

A number of studies have examined the impact of random connectivity on the ability of a sensory representation to support categorization (Babadi and Sompolinsky, 2014; Barak et al., 2013; Cortes and Vapnik, 1995; Litwin-Kumar et al., 2017; Marr, 1969; Rigotti et al., 2013). In these studies, categorization consists of dividing stimuli (odors) into two classes based on arbitrarily assigned valences. This type of categorization task emphasizes the importance of decorrelation because correlations in piriform responses have a negative impact. In our model, associations are learned for a small subset of odors and behavioral decisions are made on the basis of generalization. In this task, correlations between piriform responses to odors are essential for assessing odor similarity, and similarity provides the basis for generalization. Correlations are also essential for the consistency of generalization across individuals and between brain hemispheres despite the presence of different representations in different cortices. Lastly, in tasks in which structure in the stimulus space can be exploited, plasticity on the random layer itself can further aid the ability to categorize (Babadi and Sompolinsky, 2014; Litwin-Kumar et al., 2017; Pehlevan and Chklovskii, 2014).

Correlations between odor representations in the bulb reflect correlations between the binding of odors to receptors in the nose. These correlations are retained in the piriform despite the random connectivity from the bulb. Since individuals in a species express similar repertoires of receptors, correlations will be preserved across neuronal populations in different

piriform cortices, but the neurons responsible for these correlations will be different. In our model, individuals with different piriform cortices can nevertheless consistently assess the similarity of odors. In this manner, "a rose by any other name would smell as sweet" (Shakespeare, 2015).

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- CONTACT FOR REAGENT AND RESOURCE SHARING
- METHOD DETAILS
  - Piriform Model
  - Fly Model
  - Similar Scaling of the SNR and Two-readout Correlation Coefficient as a Function of the Number of Piriform Neurons
  - Agreement for Two Readouts
  - Agreement for a Population of Many Readouts
  - Piriform Data Analysis
- QUANTIFICATION AND STATISTICAL ANALYSIS
- DATA AND SOFTWARE AVAILABILITY

SUPPLEMENTAL INFORMATION

Supplemental Information includes two figures and can be found with this article online at https://doi.org/10.1016/j.neuron.2018.04.004.

ACKNOWLEDGMENTS

We thank Mike Shadlen and Torsten Wiesel for constructive skepticism that motivated this work, Toshi Hige and Glenn Turner for generously sharing data, Ashok Litwin-Kumar for helpful comments and suggestions, Phyllis Kisloff and Clayton Eccard for assistance in preparation of the manuscript, and Miriam Gutierrez and Adriana Nemes for general laboratory support. This work was supported by the Howard Hughes Medical Institute (453), the Gatsby Charitable Foundation, NSF NeuroNex Award DBI-1707398, and the Simons Collaboration on the Global Brain (SCGB) Award 328055. E.S.S. was supported by SCGB postdoctoral fellowship 481778.

AUTHOR CONTRIBUTIONS

Conceptualization, E.S.S., D.D.S., R.A., and L.F.A.; Software, E.S.S. and D.D.S.; Methodology, E.S.S., D.D.S., and G.B.C.; Formal Analysis, E.S.S., D.D.S., D.K., and L.F.A.; Investigation, E.S.S. and D.D.S.; Data Curation, D.D.S. and D.K.; Writing – Original Draft, E.S.S., D.D.S., R.A., and L.F.A.; Writing – Review & Editing, E.S.S., D.D.S., R.A., and L.F.A.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: October 6, 2017
Revised: February 3, 2018
Accepted: April 3, 2018
Published: April 26, 2018

REFERENCES

Aso, Y., and Rubin, G.M. (2016). Dopaminergic neurons write and update memories with cell-type-specific rules. eLife *5*, e16135.

Babadi, B., and Sompolinsky, H. (2014). Sparseness and expansion in sensory representations. Neuron *83*, 1213–1226.

Barak, O., Rigotti, M., and Fusi, S. (2013). The sparseness of mixed selectivity neurons controls the generalization-discrimination trade-off. J. Neurosci. *33*, 3844–3856.

Buck, L., and Axel, R. (1991). A novel multigene family may encode odorant receptors: a molecular basis for odor recognition. Cell *65*, 175–187.

Caron, S.J.C., Ruta, V., Abbott, L.F., and Axel, R. (2013). Random convergence of olfactory inputs in the Drosophila mushroom body. Nature *497*, 113–117.

Cho, Y., and Saul, L.K. (2009). Kernel methods for deep learning. NIPS *9*, 342–350.

Choi, G.B., Stettler, D.D., Kallman, B.R., Bhaskar, S.T., Fleischmann, A., and Axel, R. (2011). Driving opposing behaviors with ensembles of piriform neurons. Cell *146*, 1004–1015.

Cortes, C., and Vapnik, V. (1995). Support-vector networks. Mach. Learn. *20*, 273–297.

Davison, I.G., and Ehlers, M.D. (2011). Neural circuit mechanisms for pattern detection and feature combination in olfactory cortex. Neuron *70*, 82–94.

Davison, I.G., and Katz, L.C. (2007). Sparse and selective odor coding by mitral/tufted neurons in the main olfactory bulb. J. Neurosci. *27*, 2091–2101.

Eichler, K., Li, F., Litwin-Kumar, A., Park, Y., Andrade, I., Schneider-Mizell, C.M., Saumweber, T., Huser, A., Eschbach, C., Gerber, B., et al. (2017). The complete connectome of a learning and memory centre in an insect brain. Nature *548*, 175–182.

Ghosh, S., Larson, S.D., Hefzi, H., Marnoy, Z., Cutforth, T., Dokka, K., and Baldwin, K.K. (2011). Sensory maps in the olfactory cortex defined by long-range viral tracing of single neurons. Nature *472*, 217–220.

Godfrey, P.A., Malnic, B., and Buck, L.B. (2004). The mouse olfactory receptor gene family. Proc. Natl. Acad. Sci. USA *101*, 2156–2161.

Gruntman, E., and Turner, G.C. (2013). Integration of the olfactory code across dendritic claws of single mushroom body neurons. Nat. Neurosci. *16*, 1821–1829.

Hige, T., Aso, Y., Modi, M.N., Rubin, G.M., and Turner, G.C. (2015a). Heterosynaptic plasticity underlies aversive olfactory learning in Drosophila. Neuron *88*, 985–998.

Hige, T., Aso, Y., Rubin, G.M., and Turner, G.C. (2015b). Plasticity-driven individualization of olfactory coding in mushroom body output neurons. Nature *526*, 258–262.

Igarashi, K.M., and Mori, K. (2005). Spatial representation of hydrocarbon odorants in the ventrolateral zones of the rat olfactory bulb. J. Neurophysiol. *93*, 1007–1019.

Illig, K.R., and Haberly, L.B. (2003). Odor-evoked activity is spatially distributed in piriform cortex. J. Comp. Neurol. *457*, 361–373.

Iurilli, G., and Datta, S.R. (2017). Population coding in an innately relevant olfactory area. Neuron *93*, 1180–1197.e7.

Litwin-Kumar, A., Harris, K.D., Axel, R., Sompolinsky, H., and Abbott, L.F. (2017). Optimal degrees of synaptic connectivity. Neuron *93*, 1153–1164.e7.

Ma, L., Qiu, Q., Gradwohl, S., Scott, A., Yu, E.Q., Alexander, R., Wiegraebe, W., and Yu, C.R. (2012). Distributed representation of chemical features and tunotopic organization of glomeruli in the mouse olfactory bulb. Proc. Natl. Acad. Sci. USA *109*, 5481–5486.

Marin, E.C., Jefferis, G.S.X.E., Komiyama, T., Zhu, H., and Luo, L. (2002). Representation of the glomerular olfactory map in the Drosophila brain. Cell *109*, 243–255.

Marr, D. (1969). A theory of cerebellar cortex. J. Physiol. *202*, 437–470.

Meister, M., and Bonhoeffer, T. (2001). Tuning and topography in an odor map on the rat olfactory bulb. J. Neurosci. *21*, 1351–1360.

Miyamichi, K., Amat, F., Moussavi, F., Wang, C., Wickersham, I., Wall, N.R., Taniguchi, H., Tasic, B., Huang, Z.J., He, Z., et al. (2011). Cortical representations of olfactory input by trans-synaptic tracing. Nature *472*, 191–196.

Mombaerts, P., Wang, F., Dulac, C., Chao, S.K., Nemes, A., Mendelsohn, M., Edmondson, J., and Axel, R. (1996). Visualizing an olfactory sensory map. Cell *87*, 675–686.

Murthy, M., Fiete, I., and Laurent, G. (2008). Testing odor response stereotypy in the Drosophila mushroom body. Neuron *59*, 1009–1023.

Oja, E. (1982). A simplified neuron model as a principal component analyzer. J. Math. Biol. *15*, 267–273.

Papadopoulou, M., Cassenaer, S., Nowotny, T., and Laurent, G. (2011). Normalization for sparse encoding of odors by a wide-field interneuron. Science *332*, 721–725.

Pehlevan, C., and Chklovskii, D.B. (2014). A Hebbian/anti-Hebbian network derived from online non-negative matrix factorization can cluster and discover sparse features. 48th Asilomar Conference on Signals, Systems and Computers, IEEE, 769–775.

Poo, C., and Isaacson, J.S. (2009). Odor representations in olfactory cortex: "sparse" coding, global inhibition, and oscillations. Neuron *62*, 850–861.

Rennaker, R.L., Chen, C.-F.F., Ruyle, A.M., Sloan, A.M., and Wilson, D.A. (2007). Spatial and temporal distribution of odorant-evoked activity in the piriform cortex. J. Neurosci. *27*, 1534–1542.

Ressler, K.J., Sullivan, S.L., and Buck, L.B. (1993). A zonal organization of odorant receptor gene expression in the olfactory epithelium. Cell *73*, 597–609.

Ressler, K.J., Sullivan, S.L., and Buck, L.B. (1994). Information coding in the olfactory system: evidence for a stereotyped and highly organized epitope map in the olfactory bulb. Cell *79*, 1245–1255.

Rigotti, M., Barak, O., Warden, M.R., Wang, X.-J., Daw, N.D., Miller, E.K., and Fusi, S. (2013). The importance of mixed selectivity in complex cognitive tasks. Nature *497*, 585–590.

Séjourné, J., Plaçais, P.-Y., Aso, Y., Siwanowicz, I., Trannoy, S., Thoma, V., Tedjakumala, S.R., Rubin, G.M., Tchénio, P., Ito, K., et al. (2011). Mushroom body efferent neurons responsible for aversive olfactory memory retrieval in Drosophila. Nat. Neurosci. *14*, 903–910.

Shakespeare, W. (2015). Romeo and Juliet. In The Norton Shakespeare, S. Greenblatt, ed. (W.W. Norton & Co.).

Sosulski, D.L., Bloom, M.L., Cutforth, T., Axel, R., and Datta, S.R. (2011). Distinct representations of olfactory information in different cortical centres. Nature *472*, 213–216.

Soucy, E.R., Albeanu, D.F., Fantana, A.L., Murthy, V.N., and Meister, M. (2009). Precision and diversity in an odor map on the olfactory bulb. Nat. Neurosci. *12*, 210–220.

Stettler, D.D., and Axel, R. (2009). Representations of odor in the piriform cortex. Neuron *63*, 854–864.

Sugai, T., Miyazawa, T., Fukuda, M., Yoshimura, H., and Onoda, N. (2005). Odor-concentration coding in the guinea-pig piriform cortex. Neuroscience *130*, 769–781.

Takahashi, Y.K., Kurosaki, M., Hirono, S., and Mori, K. (2004). Topographic representation of odorant molecular features in the rat olfactory bulb. J. Neurophysiol. *92*, 2413–2427.

Vassar, R., Ngai, J., and Axel, R. (1993). Spatial segregation of odorant receptor expression in the mammalian olfactory epithelium. Cell *74*, 309–318.

Vassar, R., Chao, S.K., Sitcheran, R., Nuñez, J.M., Vosshall, L.B., and Axel, R. (1994). Topographic organization of sensory projections to the olfactory bulb. Cell *79*, 981–991.

Wang, J.W., Wong, A.M., Flores, J., Vosshall, L.B., and Axel, R. (2003). Two-photon calcium imaging reveals an odor-evoked map of activity in the fly brain. Cell *112*, 271–282.

Wong, A.M., Wang, J.W., and Axel, R. (2002). Spatial representation of the glomerular map in the Drosophila protocerebrum. Cell *109*, 229–241.

Zhan, C., and Luo, M. (2010). Diverse patterns of odor representation by neurons in the anterior piriform cortex of awake mice. J. Neurosci. *30*, 16662–16672.

Zhang, X., and Firestein, S. (2002). The olfactory receptor gene superfamily of the mouse. Nat. Neurosci. *5*, 124–133.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Deposited Data		
Calcium imaging data of odor responses in mouse piriform cortex	Stettler and Axel, 2009	N/A
Calcium imaging and electrophysiological data of odor responses in *Drosophila* mushroom body output neurons	Hige et al., 2015b	N/A
Software and Algorithms		
Algorithms for calculation of readout correlation and agreement	This paper	https://github.com/schafferEvan/coherentGeneralization

CONTACT FOR REAGENT AND RESOURCE SHARING

Further requests for resources should be directed to and will be fulfilled by the Lead Contact, Evan S. Schaffer (ess2129@columbia.edu).

METHOD DETAILS

Piriform Model

Our network model has feedforward connectivity and consists of three layers that we refer to as the glomerular layer, cortical layer, and readout layer, respectively, as shown in Figure 1. A fraction S_x of the N_x neurons in the glomerular layer are activated by a given odor. Glomerular representations of class odors with excess overlap f are chosen such that all such odors activate a common set of fN_xS_x glomeruli and an additional randomly selected $(1-f)N_xS_x$ glomeruli. The case with no glomeruli designated as shared $(f=0)$ corresponds to independent or nonclass odors. Response magnitudes of glomeruli whose responses are nonzero are chosen from an N_o-dimensional multivariate lognormal distribution, $\exp(\mathcal{N}(\mu_g, \Sigma_g))$, where the diagonal elements of the covariance matrix Σ_g all equal σ_g^2, and off-diagonal elements equal $\ln[f(\exp[\sigma_g^2]-1)+1]$, and N_o is the number of class odors. This yields response magnitudes whose average cross-odor correlation is f.

Each of N_y cortical neurons receives $S_{ce}N_x$ excitatory and $S_{ci}N_x$ inhibitory synaptic inputs from a random selection of glomeruli with weights, represented by matrix elements J_{ij}, set to 1 and S_{ce}/S_{ci}, respectively. The response of each cortical neuron is a threshold-linear function of the sum of its inputs, $y_i = \Theta[\sum_j J_{ij}x_j - \theta]$, where $\Theta(x)=x$ if $x>0$, and $\Theta(x)=0$ otherwise. The threshold θ is chosen such that an average of 6.2% of piriform neurons are activated by each odor, matching our imaging data. For the parameters chosen, this results in an excess of 9.2 excitatory inputs being required, on average, for a piriform neuron to be activated, i.e., $\theta \approx 9.2 * \exp[\mu_g + (\sigma_g^2/2)]$. Finally, the readout unit receives a weighted input from every piriform neuron, $z = \sum_i W_i y_i$, where $W_i = y_i^*$ after training, where the piriform response to the trained odor is y_i^*. Before training, W_i is chosen randomly with the same statistics as y. For all simulation results shown, $N_x = 1000$, $\mu_g = 0.1$, $\sigma_g = 0.5$, $S_x = 0.1$, $S_{ce} = 0.2$, $S_{ci} = 0.4$, $\theta = 11.9$.

In all quantities computed from the performance of one or multiple trained readout units, a single odor is trained, and the readout is tested with a panel of unrelated odors. Thus, readout accuracy is calculated as the fraction of novel odors correctly rejected as different from the trained odor. A more symmetric task in which the stimulus test set also includes noisy presentations of the trained odor, to which the readout should respond with the same output as the trained odor, gives qualitatively similar results.

Fly Model

Our fly model differs from our piriform model in the following parameters: $N_x = 50$, $S_x = 0.2$ (see Wang et al., 2003), $S_{ce} = 7/50$ (see Caron et al., 2013), $S_{ci} = 1$ (see Papadopoulou et al., 2011). The only other difference between these models is that in our fly model, the readout plasticity is anti-Hebbian (rather than Hebbian), so that synapses from KCs responsive to the trained odor are set to zero, and synapses from KCs not responsive to the trained odor remain strong. This feature is not essential for our results but is motivated by the experimental observation that KC-MBON synapses depress upon pairing of an odor with an unconditioned stimulus (Aso and Rubin, 2016; Hige et al., 2015a; Séjourné et al., 2011).

Similar Scaling of the SNR and Two-readout Correlation Coefficient as a Function of the Number of Piriform Neurons

As seen in Figure 2B, the signal-to-noise ratio (SNR) of a single readout and the correlation coefficient (CC) between two readouts, all trained by a Hebbian rule, have virtually identical scaling as a function of the number of piriform neurons supporting the readout. We

denote the response of the single readout to odor a by z_a and the valence of this odor by v_a. Similarly, the responses of the two readouts are denoted by $z_a^{(1)}$ and $z_a^{(2)}$. The signal-to-noise ratio for the single readout is given by

$$\mathrm{SNR} = \frac{\langle v_a z_a \rangle_a^2}{\mathrm{Var}(z_a)},$$

where the average in the numerator is over odors and the denominator is the variance across odors. The correlation coefficient between two readouts is

$$\mathrm{CC} = \frac{\left\langle z_a^{(1)} z_a^{(2)} \right\rangle_a}{\left(\mathrm{Var}\left(z_a^{(1)}\right)\mathrm{Var}\left(z_a^{(2)}\right)\right)^{\frac{1}{2}}}.$$

The numerators of these two expressions depend only weakly (as $\mathcal{O}(1) + \mathcal{O}(N_y^{-1/2})$) on the number of piriform neurons when this number is significantly greater than 1. The dependences of the denominators on this number thus determine the scaling we are discussing. Because the single readout z_a and the pair of readouts $z_a^{(1)}$ and $z_a^{(2)}$ all have the same statistics, $\mathrm{Var}(z_a^{(1)})\mathrm{Var}(z_a^{(2)}) = \mathrm{Var}(z_a)^2$ and thus the denominators of SNR and CC are identical. These two facts explain their similar scaling.

Agreement for Two Readouts

Agreement A_θ is calculated by subtracting the chance probability of agreement β from the raw fraction of similarly-responding readouts, α, and then normalizing to one, so that $A_\theta = (\alpha - \beta)/(1 - \beta)$, where $\beta = \theta^2 + (1 - \theta)^2$. The relationship between A_θ and θ can be seen more easily in the Gaussian approximation of A_θ, which makes explicit the dependence on the correlation σ,

$$A_\theta^{Gauss} = 1 - \left(\frac{1}{\pi\sqrt{1-\sigma^2}}\right)\iint_{x>\theta,\, y<\theta} dx\, dy \exp\left[\left(\frac{1}{2\sigma^2 - 2}\right)\left(x^2 - 2\sigma xy + y^2\right)\right].$$

Agreement for a Population of Many Readouts

Although the case of two readouts is of particular interest because of the application to modeling opposing hemispheres of the same brain, the agreement between more than two readouts is also of interest, with applications to readouts in multiple brain areas or in multiple individuals. We define the agreement of a population of N_z readouts, $A_\theta^{\phi,\,N_z}$, with a threshold criterion ϕ on the fraction of readouts giving the same response (either 1 or 0); in other words, a "yes" choice results when the fraction of readouts with suprathreshold values is greater than a number ϕ. Our original definition of A_θ in the two-readout case can be seen as population agreement with $\phi = 1$. As in the two-readout case, we subtract the probability of agreement occurring by chance and normalize to a maximum of 1. More precisely, we subtract the chance probability of agreement β from the raw fraction of similarly-responding readouts, α, and then normalizing to one, so that $A_\theta^{\phi,N_z} = (\alpha - \beta)/(1 - \beta)$, where β is the sum of binomial cumulative density functions, $\beta = \sum_{i=0}^{N_z(1-\phi)} (N_z/i)\theta^i(1-\theta)^{N_z-i} + \sum_{i=N_z\phi}^{N_z} (N_z/i)\theta^i(1-\theta)^{N_z-i}$. For $N_z = 2$ and $\phi = 1$, this reduces to $\beta = \theta^2 + (1-\theta)^2$.

To properly compare the required piriform resources for readout populations of varying size, we examine population agreement as a function of the total number of piriform neurons – the number of piriform inputs per readout, multiplied by the number of readout units. This is an assumption of nonoverlapping inputs to the readouts. Population agreement, A_θ^{ϕ,N_z}, for a range of ϕ values with $\theta = 0.5$ and a population of 10 readouts is qualitatively similar to the original A_θ curves (Figure S2A, left); independent of the value of the choice threshold ϕ, $A_{0.5}^{\phi,10}$ requires in excess of 10 million piriform inputs for performance to saturate. With a larger readout population ($N_z = 50$), a dependence on the value of the choice threshold ϕ emerges, such that $A_{0.5}^{\phi,50}$ looks very similar to A_θ when $\phi = 0.55$ but requires several orders of magnitude more piriform neurons when $\phi = 0.95$ (Figure S2A, right). Intuitively, the appearance of an effect of choice threshold when the readout population is large is a consequence of the ability of a large readout population to detect smaller deviations from chance: low values of the choice threshold ϕ make the population agree more often and therefore require less piriform inputs, but it also makes the chance probability of agreement higher. Only a sufficiently large readout population can distinguish between real and chance agreement and take advantage of this lower threshold. Most importantly, even in this case, performance does not saturate until the number of piriform inputs reaches ~1 million, suggesting that the importance of a large piriform cortex is general.

Piriform Data Analysis

Our analyses are based aggregated data from calcium imaging of layer 2 of piriform cortex expressing GCaMP3, GCaMP5, and Oregon Green 488 BAPTA-1 AM (Stettler and Axel, 2009). The responses of thousands of neurons to odor delivery in freely-breathing anesthetized animals were compared to assess response selectivity. Approximately 6% of piriform neurons respond selectively to odors at concentrations of 1-10 ppm. A small fraction of the neurons (less than 1%) respond to most or all presented odors. These cells may represent a separate functional class and have been excluded from further analyses. Cells responsive to a given odor are found across the piriform with no spatial preference, and the representations of different odors exhibit considerable spatial overlap, in agreement with prior work (Illig and Haberly, 2003; Iurilli and Datta, 2017; Stettler and Axel, 2009). Cells responsive to the ethologically

relevant odors TMT and mouse urine were also found to be distributed without spatial preference (Illig and Haberly, 2003; Stettler and Axel, 2009). Cells and odor responses were identified and quantified using custom software written in MATLAB as previously described (Stettler and Axel, 2009).

Nearest-neighbor statistics were calculated using the distance from each cell responsive to a specific odor within a field to the nearest other cell responsive to that odor. We then compared the distribution of these values with the distribution derived from Monte Carlo simulations in which responsive cells were drawn randomly from all the cells in the field, and nearest-neighbor statistics for each cell were then computed in the same manner. In Figure S1F, 5^{th} and 95^{th} percentiles of the Monte Carlo-derived distribution were computed separately for each distance bin.

If responsive cells were clustered, they would tend to have smaller nearest-neighbor separations than randomly distributed cells. At a representative imaging site, the *in vivo* nearest neighbor distances closely match randomly generated distances (Figure S1F). The site-averaged nearest neighbor distances in observed and randomly shuffled data are closely matched across all imaging sites and odors (Wilcoxon signed-rank test, $p > 0.3$; Figure S1G), providing no evidence for clustering. We also inspected *in vivo* piriform odor representations for local patterning by analyzing their spatial wavelengths. The spatial power spectra of *in vivo* response patterns were compared to spectra from simulations for the same sites. No significant discrepancies between the *in vivo* and randomly generated spectra are apparent for any spatial period across the sites (Figure S1H).

QUANTIFICATION AND STATISTICAL ANALYSIS

Results for readout accuracy, SNR, correlation, and agreement are averaged over input patterns and over instantiations of the network architecture.

DATA AND SOFTWARE AVAILABILITY

Software was written in MATLAB (https://www.mathworks.com/). Code used to compute quantities presented in this study is available at: https://github.com/schafferEvan/coherentGeneralization.