Cortical feedback and gating in olfactory pattern completion and separation

Gaia Tavoni\textsuperscript{a,b,1,2}, David E. Chen Kersen\textsuperscript{a,c,1}, and Vijay Balasubramanian\textsuperscript{a,b,c}

\textsuperscript{a}Computational Neuroscience Initiative; \textsuperscript{b}Department of Physics and Astronomy; \textsuperscript{c}Department of Bioengineering, University of Pennsylvania, Philadelphia, PA 19104

A central question in neuroscience is how context changes perception of sensory stimuli. In the olfactory system, for example, experiments show that task demands can drive merging and separation of cortical odor responses, which underpin olfactory generalization and discrimination. Here, we propose a simple statistical mechanism for this effect, based on unstructured feedback from the central brain to the olfactory bulb, representing the context associated with an odor, and sufficiently selective cortical gating of sensory inputs. Strikingly, the model predicts that both pattern separation and completion should increase when odors are initially more similar, an effect reported in recent experiments. The theory predicts reversals of these trends following experimental manipulations and neurological conditions such as Alzheimer’s disease that increase cortical excitability.

Pattern completion | pattern separation | olfaction | feedback | statistical modeling

Contextual information has a powerful effect on perception across a range of sensory modalities (1–8). In olfaction, experiments have demonstrated the influence of context and task demands on neural representation of odors at different levels in the olfactory pathway. In the olfactory bulb (OB), where odor information is first processed before passing to cortex, context-dependent changes in both single-neuron and collective bulb activity have been observed during and following learning (9–17). Context can also reshape the representations of odors in the olfactory cortex: when odors are associated with the same or different contexts, the corresponding cortical activity undergoes pattern completion (increased response similarity) or separation (decreased response similarity) respectively (18, 19). The mechanisms underlying such context-induced transformations are of great interest in sensory neuroscience.

The OB and cortex are notably coupled to one another. Mitral cells (MCs) and tufted cells (TCs) from the bulb project to several higher brain regions, including piriform cortex (PC), anterior olfactory nucleus (AON), olfactory tubercle, entorhinal cortex, and amygdala (20). In particular, experiments have highlighted that the anterior PC is activated by convergent and synchronous inputs from the bulb: coincident activation of a few glomeruli within a short time window (21) is required to induce spiking in cortical pyramidal neurons, a mechanism that is thought to be important for decoding complex combinations of chemical features (22, 23). In turn, the bulb receives extensive centrifugal feedback from several areas of the central brain, including PC, AON, midbrain, amygdala, hippocampus, and entorhinal cortex (24–41). This feedback predominantly targets granule cells (GCs) and other bulb interneurons, enhancing or suppressing their activity (24, 26, 42–47), but may also directly excite the MC/TCs (29, 30, 43). These reciprocal interactions are known to play important roles in forging odor-context associations (10, 14, 15, 41, 48–52) as well as in generating beta oscillations (53–55), which are thought to reflect coordination between the OB and cortical areas during learning. However, the precise mechanism by which centrifugal feedback to the OB can effect change in cortical odor representation remains unclear.

Interestingly, the structural organization of these feedforward and feedback projections is highly disordered. MC/TCs project to the cortex in an apparently random fashion (22, 56–58), while centrifugal feedback fibers are distributed diffusely over the OB without any discernible spatial segregation (25, 30). We hypothesized that these diffuse feedback signals carry unstructured representations of context, which modulate odor responses in the OB and in turn entrain robust pattern completion and separation in piriform cortex (PC). To test this hypothesis, we constructed a statistical, analytically tractable model of the OB and its projections to the PC, which we further extended by incorporating an anatomically-faithful network of interactions between OB excitatory and inhibitory cell types along with a realistic distribution of projections from the OB to the PC. Under minimal assumptions about the statistics of odor and feedback inputs, we show that changes in bulb firing rates driven by unstructured feedback lead to pattern separation and completion in cortex. The model predicts that both the pattern separation and, counterintuitively, pattern completion, should increase with initial odor similarity. This prediction matches results from recent experiments (18, 19, 59). Our results require that cortical units selectively gate coincident inputs as PC pyramidal neurons in fact do (21, 23). The model also predicts that increases in cortical excitability, either by experimental manipulation, or in neurological conditions such as Alzheimer’s disease, will alter or even reverse these trends in cortical pattern separation and completion, and hence induce characteristic changes in odor perception and behavior.

Results

A statistical mechanism for associating contexts to odors. We hypothesize that unstructured centrifugal feedback to a more peripheral brain area (here, the olfactory bulb (OB)) can trigger changes in responses to odor inputs that lead to convergence or divergence of activity patterns in a more central brain area to which the peripheral area projects (here, the piriform cortex (PC)). This convergence or divergence can underpin generalization and discrimination in sensory behavior (Fig. 1A). To test whether and how this mechanism may

G.T. conceived the study and performed research on the statistical model; D.K. performed research on the mechanistic model; all the authors discussed every aspect of the research and wrote the paper. The authors have no competing interests.

\textsuperscript{1}These authors contributed equally to this work.

\textsuperscript{2}To whom correspondence should be addressed. E-mail: tavoni@sas.upenn.edu
work, and to derive its predictions, we first construct a simple statistical model of the interactions in the olfactory system.

**Structure of the model.** Consistent with basic anatomical features of the olfactory system (Fig. 1B), our model contains a sensory layer and a cortical layer, the OB and the PC respectively. Sensory projection neurons in the first layer, here taken to be mitral cells (MCs) in the OB, are grouped into “modules”, where each module comprises the set of MCs that project to the same pyramidal neuron in the cortical (PC) layer. The sensory (OB) layer is thus composed of $N$ modules, and the cortical (PC) layer of $N$ corresponding units that represent pyramidal neurons receiving direct inputs from the modules.

**Modeling odor inputs.** The response of an OB module $q$ to an odor input is summarized by the change in the total firing $R_q$ of its constituent MCs compared to baseline. Thus the response for the bulb as a whole is described by the module firing rate vector

$$ R = (R_1, R_2, \ldots, R_N) \ . \quad [1] $$

Neural responses fluctuate, so we consider module $i$ to be odor responsive only if $R_i > \theta_i$, where $\theta_i$ is a threshold for statistically significant activity. Experiments have indicated that piriform neurons only respond if sufficiently many of their inputs are active within a short window ($21, 22$). Thus, in our model, a cortical neuron receiving input from module $q$ is denoted as “active” only when $R_q > \theta_q$, where $\theta_q$ quantifies a threshold for cortical activation. In the normal brain, the cortical threshold is significantly higher than the threshold for a module to be considered odor responsive ($\theta_m \gg \theta_n$) ($21, 22$).

**Modeling context-induced feedback inputs.** We model context as the effect produced by cortical feedback inputs to the OB. In the OB, such feedback can lead to increases or decreases in MC firing and consequently the strength of inputs to PC, represented in our model by the module firing rates. Centrifugal excitation of interneurons such as GCs and periglomerular cells reduces MC firing ($24, 25, 27–30, 43$), while other modes of feedback, such as direct excitation of MCs ($29, 30$), neuromodulation of MC excitability ($32, 34, 37, 38, 40$), and excitation of deep short axon cells (which drives feedforward inhibition of GCs) ($24, 26, 42$) enhance MC activity. As a result, we represent the effect of feedback as a change in module firing rates:

$$ \Delta R = (\Delta R_1, \Delta R_2, \ldots, \Delta R_N) \quad [2] $$

Where $\Delta R_i < 0$ and $\Delta R_i > 0$ respectively represent decreases and increases in the firing rate of module $i$. 

---

**Computation of the cortical responses to odor and feedback inputs**

![Fig. 1. A statistical mechanism for context-induced pattern completion and separation in the olfactory system.](image)

(A) Context is information present in the environment that is salient to behaviour, e.g. the location of a reward associated with an odor. Left: in two-alternative forced-choice tasks, when distinct odors are associated with the same context (e.g. a left reward port), animals learn to generalize their choice (approaching the left port) across odors; after training, the neuronal ensemble responses to the odors in the PC are more correlated (pattern completion, (18)). Right: when odors are associated with alternative contexts (e.g. opposite reward ports), animals learn to discriminate the stimuli and make different, odor-specific, choices (approaching the left versus right port); after training, the cortical ensemble responses are less correlated (pattern separation, (18)). (B) Sketch of the olfactory system and statistical model. In the olfactory bulb, mitral cells (MCs) process olfactory inputs in structures called glomeruli, then broadcast the sensory information to piriform cortex (PC) and other areas. Projections to the PC are seemingly random and glomerulus-independent. The main bulb interneurons, the granule cells (GCs), implement lateral inhibition between the MCs. Both MCs and GCs receive centrifugal feedback from the central brain, conveying information about context. In the statistical model, the olfactory bulb is represented as an ensemble of “modules” (red circles), with each module defined as the set of MCs that project to the same cortical neuron, and the activation probability of cortical cells is a threshold function of the modules’ responses. (C) A geometrical interpretation of Eqs. 4, 5, 6. Each square represents the space of module responses to two odors (A and B). Left: the number of cortical cells activated by odor A ($C_A$ in Eq. (4)) is equal to the number of modules with responses higher than the cortical activation threshold $\theta_i$ (green rectangle); likewise for $C_B$ (red rectangle) and $C_{AB}$ (overlap between the green and red rectangles). Middle: under assumptions of uniformly distributed module responses (see text), $C_A$, $C_B$ and $C_{AB}$ are fractions of the numbers of modules with significant responses (i.e. higher than the module response threshold $\theta_m$) to, respectively, odor A (i.e. $N_A$, green rectangle), odor B (i.e. $N_B$, red rectangle) and both (i.e. $N_{AB}$, overlap between the two rectangles). Each fraction is the ratio between areas of the same color in the left and middle panels, yielding factor $Q_{ij}$ in Eq. 5. Right: stochastic feedback representing contextual changes in the distribution of module responses. Modules with responses to either odor within $\Delta R$ from the cortical activation threshold and targeted by positive/negative feedback (with probabilities $p_+$ and $p_-$, respectively) are brought above/below the cortical threshold, changing correlation between the cortical odor representations (Eq. 6).
Additionally, we assume that the correlation between the changes in modules’ responses induced by any two given feedback inputs directly reflects the similarity between the contexts that elicit those feedback inputs. We quantify this correlation as the normalized overlap (i.e. cosine of the angle) between the vectors of response changes:

\[ \rho_{FB} = \frac{\Delta R(F_1)}{\|\Delta R(F_1)\|} \cdot \frac{\Delta R(F_2)}{\|\Delta R(F_2)\|} \]  

where F1 and F2 denote the two feedback inputs. In the most extreme case, feedback inputs that induce identical changes in the modules’ responses represent identical contexts, e.g. two odors being associated with a reward at the same location (Fig. 1A, left). Conversely, those that induce anticorrelated changes represent antithetical contexts, e.g. two odors being associated with a reward at different locations (Fig. 1A, right).

**Defining correlation.** The correlation between cortical responses to two odors A and B before feedback is quantified by a normalized overlap of the cortically activated modules:

\[ \rho = \frac{C_{AB}}{\sqrt{C_A C_B}} \]

where \( C_{AB} \) is the number of cortical units activated (or, equivalently, modules in the bulb that are driven to a firing rate greater than \( \theta_c \)) by both odors, and \( C_A \) and \( C_B \) are the numbers of cortical units activated by odors A and B respectively.

**Change in correlation between odor representations following feedback varies linearly with the initial similarity.** Feedback-induced changes in OB module responses impact the activation state of cortical units, thus changing the correlation between cortical responses to odors. We analyze these changes first in a simple analytical model where the effects and their conceptual origin are qualitatively apparent, and then in biophysically realistic scenarios where numerical analysis is necessary.

To capture the overall features of the system while simplifying for analytical tractability we first suppose that: (a) the module firing rate distribution is uniform below and above the cortical threshold for responding (i.e. whose firing rates are also above \( \theta_c \)), (b) the firing rate of any module that responds to the odors can be obtained by replacing \( \theta \) with \( \theta - \Delta R \) (for both odors). Therefore the final correlation in the cortical activation state of cortical units, thus changing the correlation between the feedback inputs. This linear relationship can be understood intuitively as follows: feedback changes correlation by pushing some modules above or below the cortical threshold (depending on the feedback sign); these activated/deactivated modules are typically close to the cortical threshold and odor-responsive before feedback (Fig. 1C, right). Thus, they can be expressed as a fraction of the total number of odor-responsive modules, similarly to Eq. (5). This gives rise to the \( Q \) term in Eq. 6. If feedback is sufficiently strong compared to the difference between the cortical and response thresholds, it may also activate some modules that are not odor responsive. This effect gives rise to the constant term \( Q_1 \).

**Identical contexts induce pattern completion that increases with the initial odor similarity.** Suppose first that the contexts associated with two odors A and B are identical and, through feedback, increase the responses of some modules while decreasing the responses of some other modules by the same amount. As a simple illustration, consider the special case that all modules receive the same positive feedback \( \Delta R \). In this case, all modules with responses to odor A or B that are close enough to the cortical threshold \( \theta_c - \Delta R < R_c < \theta_c \) are brought above threshold by the feedback (Fig. 2A, left); as a result, the cortical neurons to which they project are activated by the feedback. These context-induced changes are equivalent to an effective decrease of the cortical activation threshold by \( \Delta R \) (for both odors). Therefore the final correlation in the cortical responses to the odors can be obtained by replacing \( \theta_c \) with \( \theta_c - \Delta R \) in Eq. 5, and is linear in the initial correlation:

\[ \rho_f = \frac{\Delta R - \theta_c + \Delta R}{R_{\max} - \theta_m} \frac{N_{AB}}{\sqrt{N_A N_B}} \]

In this scenario, the context-induced change in correlation is

\[ \Delta \rho = \frac{\Delta R}{R_{\max} - \theta_m} \rho_i \]

which is a special case of Eq. (6) with \( Q_1 = 0 \) and \( Q_2 > 0 \). Thus, the correlation in the cortical responses to odors A and B increases in proportion to the initial correlation \( Q_2 > 0 \) in Eq. 6 when the odors are presented in identical contexts eliciting the same excitatory feedback (Fig. 2A, middle).

More generally, suppose that feedback increases or decreases the firing rates of some modules by an amount \( \Delta R \) in response to either odor. In this case, feedback-receiving modules with responses that lie close enough to the cortical threshold for
either odor may cross this threshold, either becoming active or inactive (Fig. 2A, right). Pictorially, modules receiving positive feedback in the gray region of Fig. 2A, right (i.e. just below the threshold for either odor) now are active for both odors, adding to the cortical correlation. Meanwhile, modules receiving negative feedback in the brown region of Fig. 2A, right (i.e. just above the threshold for both odors) will become inactive for at least one odor, and thus stop contributing to the cortical correlation. Note that the gray region is bigger than the brown region, i.e., there are initially more modules activated by just one or no odor rather than both. Thus, provided that the fraction of modules receiving positive feedback is not too much smaller than the fraction of modules receiving negative feedback, the net effect will be to increase the relative fraction of modules active for both odors, i.e., the cortical overlap: pictorially, more modules move from the gray to brown region than vice versa because there are more of them in the first place. In other words, identical feedback to the bulb leads to pattern completion in the cortical responses. Quantitatively, the normalization of the correlation by the geometric mean of the modules active for each odor (Eq. 4) has the effect that correlated feedback increases cortical correlation even if the feedback suppresses the firing rate of most (but not all) modules (details in Supplementary Information). Furthermore, in realistic conditions, only odor-responsive modules (above $\theta_m$) can be pushed by feedback above the activation threshold $\theta_c$; thus the term $Q_1$ in Eq. 6 vanishes (details in Supplementary Information). As a result, pattern completion increases proportionally to the initial odor similarity.

**Opposite contexts induce pattern separation that increases with the initial odor similarity.** Next, consider a scenario where the contexts associated with odors A and B are antithetical, and through feedback, cause the responses of some modules to change oppositely for the two odors.

As a simple illustration, consider a situation where the context for odor A increases the firing rate of all modules by $\Delta R$, while the context for odor B decreases the firing rate of all modules by the same amount. By an argument similar to the one above for the case of identical feedback, these effects are equivalent to an effective decrease/increase of the cortical...
threshold $\theta_c$ by $\Delta R$ for the two contexts (Fig. 2B, left). Thus, \[
\rho_f = \frac{\sqrt{(R_{\text{max}} - \theta_c + \Delta R)(R_{\text{max}} - \theta_c - \Delta R)}}{R_{\text{max}} - \theta_m} \frac{N_{A,B}}{\sqrt{N_A N_B}} [9]
\]
and \[
\Delta \rho = \left( \frac{\sqrt{(R_{\text{max}} - \theta_c + \Delta R^2)}}{R_{\text{max}} - \theta_c} - 1 \right) \rho_0 . [10]
\]
This is a special case of Eq. (6) with $Q_1 = 0$ and $Q_2 < 0$. Thus the cortical correlation decreases in proportion to the initial response similarity ($Q_2 < 0$ in Eq. 6, Fig. 2B, middle).

More generally, suppose that the feedback inputs to some modules cause their firing rates to change oppositely for the two odors. If a module was active for both odors, but close to the cortical threshold (brown region in Fig. 2B, right), opposite feedback will typically make it inactive for one odor, decreasing the cortical correlation. Conversely, only modules that are just below the cortical threshold for one odor and sufficiently above the cortical threshold for the other odor (gray regions in Fig. 2B, right) can become active for both odors as a result of opposite feedback. Because the gray regions are smaller than the brown region, the net effect of opposite feedback is a decrease in the cortical correlation. As in the case of identical contexts, in realistic conditions the term $Q_1$ in Eq. 6 vanishes (details in Supplementary Information); thus, the amplitude of pattern separation increases proportionally to the initial odor similarity.

**Pattern completion vs. separation for general feedback conditions.**

We tested that the qualitative conclusions above persist across diverse conditions where the feedback strength ($\Delta R$), the fractions of modules receiving positive and negative feedback, and the levels of correlation between the feedback inputs vary (Fig. 3A–E). Feedback in our model is minimally constrained in the sense that it is not designed to target specific MCs, e.g. based on their odor tuning properties or locations in the bulb.

We start with a scenario in which $\sim50\%$ of the modules are affected by feedback, and the majority of the feedback ($\sim75\%$) is negative for one odor, i.e., suppresses firing rates. If we assume that the feedback inputs for the two odors target random (independent) sets of modules, $50\%$ of the modules that are affected by feedback for the first odor will also be affected, on average, by feedback for the second odor. We also take the feedback strength $\Delta R$ to be $\sim20\%$ of the maximum firing rate. Fig. 3A shows that under these conditions, if we have a high cortical threshold $\theta_c$ so that about $10\%$ of the cortical units are activated by each odor as in the brain (23), then highly correlated contexts induce pattern completion, while uncorrelated or anticorrelated contexts lead to pattern separation. The pattern completion and separation both increase with the degree of initial correlation in the odor responses ($Q_2 > 0$ and $Q_2 < 0$ respectively in Eq. 6).

Figs. 3B–E demonstrate that the same result persists across: (i) different feedback strengths, and over larger ranges of cortical thresholds for stronger feedback, (ii) different fractions of modules targeted by feedback, (iii) different fractions of negative versus positive feedback, and (iv) different overlaps between the sets of modules targeted by feedback for the two odors. If the fraction of negative feedback is low, the transition between pattern separation and pattern completion at some intermediate value of the feedback correlation occurs even for low cortical thresholds (Fig. 3D, left). However, in the brain,
feedback to the bulb is thought to be dominated by projections to the inhibitory granule cells and periglomerular cells (24, 25, 27–30, 43), although there are also excitatory feedback to mitral cells (29, 30) and neuromodulatory mechanisms that increase excitability (32, 34, 37, 38, 40).

Finally, we noted that the firing rate of modules should be determined by the sum of the firing rates of the component mitral cells. Thus, by the central limit theorem, the module firing rates should be normally distributed. Likewise, feedback affects individual mitral cells, and the resulting change in the firing rate of a module is determined by the sum of these changes for the component mitral cells. Thus, the effect of feedback on the module firing rates should also be normally distributed. We included these normal distributions numerically, since an analytical treatment was not possible. We also replaced the hard threshold for cortical activation in the analytical model with a sigmoid acting on the olfactory bulb module responses with a soft threshold at $\theta_c$. We then simulated $N = 10,000$ cortical units receiving inputs from the corresponding number of bulb modules, and responding to two odors A and B. Before feedback, we took the responses of bulb modules to be normally distributed for each odor. We controlled initial odor similarity by changing the fraction of modules that responded to both odors as opposed to only one. We then modeled the effects of contextual feedback by adding normally distributed firing rate changes to the modules. The mean and standard deviation of the distributions were picked so that the range of module responses would be comparable to the analytical analysis. We tested that the results are robust to broad changes in these parameters (Fig. S1). The normalized overlap between the cortical responses to the odors measures the initial correlation. Pattern completion or separation occur if the the overlap after feedback is greater or less than the initial overlap.

Fig. 3F shows the results obtained for a large number of randomly generated odor pairs with different initial similarities and different feedback parameters. The results robustly recapitulate the analytical findings in the simplified model: when the cortical threshold is realistically high, correlated and anticorrelated feedback effects on the OB modules induce, respectively, pattern completion and pattern separation in cortex, both increasing linearly with the odor initial similarity (Fig. 3F).

Summary. Overall, these results indicate that, in our model, if the cortical threshold is high as in the brain, sufficiently similar or different contexts will robustly induce pattern completion or separation in cortical responses. What is more, both pattern completion and separation increase with initial odor similarity. Precisely these trends are seen in experiments (18, 19, 59), along with high activation thresholds for pyramidal neurons in the piriform cortex (21, 23), and diverse forms of unstructured excitatory and inhibitory feedback to the olfactory bulb.

Mechanistic model. Above, to facilitate analysis, we described activity in the bulb and the effects of feedback in terms of independent firing rates, and changes of firing rates, of bulb modules. Mechanistically, these modules comprise overlapping groups of mitral cells, and feedback targets individual mitral cells and the granule cells that inhibit them. The granule cells are in turn connected in a network, so the precise effect they have on mitral cells during combined odor driven and

---

Fig. 4. A mechanistic model produces trends in pattern completion and separation similar to those predicted by the statistical framework (A) The probability of connection is a function of the overlap of the mitral cell and granule cell dendritic trees, here represented as a disc and cone, respectively. Different parameters in the mean-field distributions of these dendritic trees produce different probability curves for any given MC-GC pair. (B) The algorithm in (A) is used to generate a network of MCs interconnected by GCs. Network dynamics are simulated via the Izhikevich model (60), with each neuron in the network receiving external inputs in the form of odor or feedback as well as reciprocal inputs resulting from the connectivity of the network. Odor input is modeled as an oscillatory current to mimic the respiratory cycle while external feedback is modeled via a constant current input. (C) For high cortical thresholds, the normalized overlap between odor representations increases (pattern completion) or decreases (pattern separation) linearly in the initial overlap, according to the correlation in the feedback inputs. In all cases, we simulated 10,000 MCs grouped into 500 glomeruli and 100,000 cortical cells, each sampling 7% of the MCs, with odor targeting 10% of the glomeruli, positive feedback targeting 20% of the MCs, and negative feedback targeting 85% of the MCs.
feedback activation is difficult to anticipate. To incorporate these considerations we refined our analysis with a detailed mechanistic model employing a biophysically realistic network in the bulb (Fig. 4A,B).

Briefly, to create this network, we modeled each cell in terms of its dendritic tree, since interactions between the two cell types are dendrodendritic (44, 61–67); roughly mimicking the anatomy, MC trees were laid in laminar discs and GC trees projected up in inverted cones (68–71). These cells were embedded in a 3-dimensional space representing the layers of the OB. The x-y location of the center of a MC dendritic disc depended on the location of its respective glomerulus (which was in turn placed randomly) (72, 73), while its z-position depended on whether it was a Type I or Type II MC (68–70). The GCs in our model were deep-type (71) (since MCs primarily interact with deep GCs), so while the x-y location of the vertex of a GC tree was random, its z-position was confined to roughly the lower half of the model space. The probability of a connection between any given MC and GC was then a function of the geometric overlap between their dendritic trees (Fig. 4A, details in Supplementary Information).

To simulate network dynamics, MCs and GCs were modeled as Izhikevich neurons (60), with parameters selected to match known electrophysiological data (70, 71, 74). Odor input was represented as an oscillatory current into the MCs, while positive and negative feedback to the system were represented as pulses of constant excitatory current to a random subset of the MCs and GCs respectively. We determined that a sufficiently large number of MCs (>5000) was required to achieve statistically stable results (Fig. S4A). The large size (especially given that there are around an order of magnitude more GCs than MCs (75)) meant that computational limitations precluded us from directly performing the spiking simulation for all the different odor and feedback tests. We circumvented this limitation by using the spiking model to extrapolate distributions of firing rates in MCs following odor input and reciprocal feedback from GCs, both with and without external feedback to MCs and GCs (see Supplementary Information, Fig. S2). This allowed us to simulate the detailed dynamics of larger numbers of MCs efficiently by summarizing the effects of the granule cell network in terms of the distribution of effects on mitral cell firing.

To simulate cortical responses, we allowed K cells to sample randomly from a fraction q of the M MCs, with q ~ 0.07, consistent with experimental measurements (56). Each group of sampled MCs thus constituted a module in the terminology used previously. MCs were also partitioned into G glomerular, non-overlapping sets, each representing the set of MCs associated to the same glomerulus. We modeled the response to an odor as the evoked firing rate over a single sniff (a length of time sufficient for a rodent to distinguish between odors (76)) in a fraction $f_{odor}$ of the glomeruli. Thus, the MCs in the $f_{odor}$ G glomeruli had a non-zero probability of having a firing rate greater than 0, while the remaining MCs had firing rates equal to 0, after subtracting baseline firing. For the fraction of MCs that were selected to receive odor input, the firing rate for each MC was drawn from a distribution fitted to data from the detailed spiking model of the bulb.

We determined the input to each cortical unit from the sum of the firing rates of MCs that projected to it. To account for the effects of cortical balancing (77), we subtracted the mean cortical input from the input to each unit, and then passed the result through a nonlinear activation function (a sigmoid). This ensured that the strongly activated cortical units tended to receive the highest rate, and hence most coincident, inputs. This sequence of steps yielded a vector of odor-induced firing rates over a sniff as the cortical representation of the odor.

Finally, we modeled feedback as a vector of firing rate changes in a fraction $f_{FB}$ of randomly selected mitral cells. These changes could be induced either by direct feedback to the mitral cells (29, 43) or indirect inhibitory feedback through the granule cells (24, 27, 29, 30, 42, 78), all of which have the net effect of modifying MC firing rates. The specific pattern of feedback, i.e. the affected MCs and whether the net effect on each increased or decreased the firing rates, effectively defined the context associated to the odor presentation.

Of note, we determined through simulations in the spiking network that negative feedback patterns to the MCs arising indirectly from positive feedback to the GCs are necessarily diffuse and unstructured. Specifically, we found that targeting non-overlapping sets of about the same number of GCs nonetheless produced highly correlated changes in the MC firing rates because the pattern of inhibition is largely determined by the structure of the GC-MC network (Fig. S3). Thus, the identity of the MCs which are affected by negative feedback through the GCs, and the degree to which they are affected, generally depend only on the fraction of GCs targeted rather than which GCs are targeted in particular. As a consequence, in our mechanistic model, the fraction of targeted GCs is the primary determinant of which MCs receive inhibition and how much their firing rate is suppressed.

An in silico experiment on pattern completion vs. pattern separation.

We tested how similarities and differences in feedback affected the cortical representation of odors in our mechanistic model. To this end, we generated pairs of odors, with each odor targeting $f_{odor}$ G glomeruli, where the two odors shared different fractions of targeted glomeruli (Fig. 4B). We then computed the normalized overlap $\rho$ between cortical responses to pairs of odors before and after addition of feedback for different feedback regimes as the cosine of the angle between the cortical firing rate vectors (which is equivalent to Eq. 4 in the analytical model), in parallel with the measure of feedback correlation in Eq. 3.

We first considered the situation where the threshold of the sigmoidal activation function is high leading to sparse firing in the cortex as seen in experiments (23, 79–81). We presented pairs of odors with different degrees of initial correlation and different amounts of feedback correlation (as quantified in Eq. 3). We found that strongly correlated feedback led to pattern completion (increased overlap in cortical responses). Likewise anticorrelated feedback led to pattern separation (decreased overlap in cortical responses). Notably the pattern completion and separation both increased linearly with the initial overlap of cortical responses (Fig. 4C).

All told, these results recapitulated the predictions of the statistical model in a detailed mechanistic setting. One difference from the abstract analysis is that in this mechanistic model the only form of direct negative feedback goes through the granule cell network and is thus non-specific to particular MCs. As a result, strongly anticorrelated feedback is hard to achieve, but may be possible in the brain through targeted neuromodulatory effects or through feedback that suppresses
granule cells. In addition, because the mitral cells are embedded in a granule cell network, excitatory feedback to MCs necessarily induces some inhibitory feedback disynaptically through the granule cells. Thus, in this realistic mechanistic model there are constraints on the achievable forms of positive and negative feedback. Stronger feedback anticorrelation would further enhance pattern separation.

The necessity of a multi-layer architecture. These results depend critically on a two-layer architecture where activity in the bulb is modulated by feedback and then passed through a nonlinearity with a high threshold in the cortex. To see this, we can remove the cortical layer from our statistical model and study how the correlation in bulb responses to odors changes due to feedback. As above, we assume that these responses prior to feedback and the feedback-induced changes are both normally distributed. We also maintain the same feedback correlations and statistics as in Fig. 3F. We then compute the correlation (normalized overlap) between bulb responses to odors A and B, before and after adding feedback. Fig. 5A shows that the effects induced by feedback in the bulb are substantially different from those that arise in the cortical layer.

In particular, pattern completion can be achieved only when the feedback correlation is very high, and even then decreases with increasing odor similarity, oppositely to the trend in the cortical layer. A single-layer model is equivalent to a model with multiple layers paired by linear transfer functions. Thus, the form of pattern completion of Fig. 3F can only emerge in a network with at least two layers, linked by a non-linear (high-threshold) transfer function, and feedback targeting the input layer.

A prediction: partial reversal of effects when cortical excitability is increased. In the normal brain, pyramidal cells in the piriform cortex have a high threshold for activation and respond only when there is coincident input from multiple mitral cells (23, 82). However, both experimental manipulation and neurological conditions such as Alzheimer’s disease can cause increases in cortical excitability, or, equivalently lower the cortical activation threshold. Our model makes striking predictions for these conditions that can be tested.

First, Figs. 3A–E show that when feedback is predominantly inhibitory, as it is in the brain, pattern completion only arises by our proposed mechanism within a range of cortical activa-
tion thresholds that are high, and include the typical values expected for pyramidal neurons to achieve realistic levels of cortical activation of about ~10% for any odor (23). Thus we predict that increased excitability in the piriform cortex (or, equivalently, reduced activation thresholds) will impair pattern completion and thus the behavioral ability to generalize. Specifically, if cortical excitability increases moderately, our model predicts that any feedback inputs (anticorrelated as well as correlated) will induce a weak pattern separation effect, increasing with initial odor similarity (light blue regions in between the black and gray dashed lines in Fig. 3A–E). If the increase in cortical excitability is sufficiently large, our model predicts more complex effects (green regions in Fig. 3A–E), as explained below. Qualitatively, in the high-threshold regime, only modules q that already have a significant response $R_q$ to an odor input ($R_q > \theta_c - \Delta R > \theta_m$) can exceed the cortical threshold due to feedback, whereas if the threshold is sufficiently low ($\theta_c \lesssim \theta_m + \Delta R$), feedback can push some modules that were not initially odor responsive to be above the cortical threshold. This can influence the final cortical correlation and change results qualitatively (Fig. 5B–D, Methods and Supplementary Information). Fig. 5B–C show numerical and analytical results from our statistical model at a low cortical threshold. In particular, Fig. 5C shows our analytical results with identical (top) and opposite (bottom) feedback. We see that some of the effects at high threshold are reversed: (a) when the initial cortical correlation is high, similar contextual feedback to the bulb can lead to cortical pattern separation instead of pattern completion ($\Delta \rho$ becomes negative in Fig. 5C, top right), (b) when the initial cortical correlation is low, dissimilar contextual feedback to the bulb can lead to cortical pattern completion instead of pattern separation ($\Delta \rho$ becomes positive in Fig. 5C, bottom right), and (c) while pattern separation for dissimilar contexts increases with increasing initial odor similarity, pattern completion for similar contexts decreases with increasing odor similarity (negative slope in Fig. 5C, top right), reversing the trend at high threshold. These reversals at low threshold are confirmed by analysis of more general feedback conditions and response distributions (green regions in Fig. 3A–E, Fig. 5B and Supplementary Information). To confirm the predictions in the mechanistic model, we considered a situation where the threshold of the sigmoidal activation function is low, leading to greater cortical excitability for the particular sensory inputs: that is, feedback is not designed to target specific sensory neurons (e.g. based on their tuning properties). Effective forms of pattern completion/separation arise statistically via a mechanism that simply adjusts the level of correlation in otherwise unstructured feedback patterns to reflect the similarity between the contexts. Feedback signals with these simple properties can be implemented in many ways and diverse sensory areas of the brain, making this mechanism applicable across modalities.

Discussion

We have identified a simple, robust mechanism whereby unstructured contextual feedback from the central brain to peripheral sensory areas can engender pattern completion and separation in cortex, with, presumably, associated effects on the perception of odors. A striking feature of our mechanism is that the feedback does not have to specifically target neurons that contribute to cortical activity for the particular sensory stimuli that are being contextually associated or dissociated. Rather, we find that random centrifugal projections producing mixed inhibitory and excitatory effects will suffice. In this mechanism, correlated feedback broadly leads to pattern completion, while uncorrelated or anticorrelated feedback to pattern separation. The indirect routing of the feedback to a sensory area that has convergent, strongly gated, projections to the cortex has the striking effect that the strength of pattern completion and separation are predicted to both increase linearly with the initial odor similarity. If the gating were weaker, i.e., if the threshold for cortical activation were lower leading to less sparse activity, some of these effects would reverse. We demonstrated that our results are robust to broad changes in the statistics of odor inputs and feedback patterns, and that they are specific to networks with at least two layers paired by a non-linear, high-threshold input-output function, as seen in cortex. Finally, we showed that this mechanism is realized in a detailed mechanistic model of the circuit architecture of the early olfactory system, with mitral cells in the olfactory bulb projecting to piriform cortex, and in turn receiving centrifugal feedback both directly and indirectly from the granule cell network of the bulb.

Generality of the proposed mechanism. Although our work focused on the olfactory system, similar mechanisms could be implemented in other areas of the brain and support pattern completion/separation for other sensory modalities. Indeed, all sensory cortices satisfy the key structural requirements of our statistical model: (1) a multi-layer architecture; (2) convergent projections from one layer to the next and increasingly selective gating of the sensory inputs (as demonstrated by a reduction in neural activation levels across the sensory hierarchy (83)); and (3) the presence of feedback carrying information about the context of the sensory stimuli towards the lower sensory layers. We showed that pattern completion and separation effects that emerge in neuronal networks with these features rely only on minimally constrained feedback signals that are agnostic to the specific neuronal activation patterns induced by the sensory inputs: that is, feedback is not designed to target specific sensory neurons (e.g. based on their tuning properties). Effective forms of pattern completion/separation arise statistically via a mechanism that simply adjusts the level of correlation in otherwise unstructured feedback patterns to reflect the similarity between the contexts. Feedback signals with these simple properties can be implemented in many ways and diverse sensory areas of the brain, making this mechanism applicable across modalities.

For the olfactory system in particular, we showed explicitly that direct excitatory feedback to the MCs and indirect, GC-mediated inhibitory feedback can yield cortical pattern completion or separation depending on the correlation between the induced responses in the bulb, and these effects increase linearly with the initial odor similarity. Interestingly, while pattern separation can be achieved with 100% inhibitory feedback, pattern completion can only arise in presence of some excitatory feedback, or a combination of excitatory and inhibitory feedback (Supplementary Information). These results suggest an etiology for the diversity of feedback modes in the olfactory system and the brain at large.

Model predictions and experimental tests. Key behavioral predictions of our theory follow from the assumption that similarities and differences in the perception of an odor are directly related to the similarities and differences in their cortical representation. Thus, our theory suggests that perceptual learning should be strongly affected by cortical excitability, and, equivalently, the response threshold of cortical neurons.
At low excitability (or high threshold), which is the normal state of sensory cortex, odors presented in related contexts should be perceived as more similar than they are, and the extent of this perceptual change should increase with the initial similarity. Thus the more similar the odors smell initially, the more identical they should seem after presentation in the same context. By contrast, if excitability is high (or threshold is low), the perceptual change in response to presentation in related contexts should be greatest for odors that were initially dissimilar. Thus, if cortex is highly excitable, in generalization tasks stimuli that are very different should be disproportionately perceived as similar and potentially confused with one another, while perceptual association of initially similar stimuli would be less effective or not achieved at all. By contrast, we predict that, at both low and high thresholds, the ability to discriminate odors associated with different contexts will increase with initial odor similarity.

Some of these predictions are already supported by experimental evidence from studies conducted in rodents. In particular, both (18) and (19) found that association of odor mixtures with different contexts correlated the cortical neural responses to the mixtures (pattern separation) only in difficult discrimination tasks, i.e. when the odor mixtures were perceptually similar. This experimental finding is consistent with our prediction of a positive gain for pattern separation as a function of the initial odor similarity. Furthermore, both studies reported an increase in cortical response correlation (pattern completion) when odors were associated with the same context (the same reward port in (18), the same positive valence in (19)). In support of our model predictions, the perceptual effects measured in (18) also strongly suggest a positive gain for pattern completion.

Overall, the predictions following from our proposed mechanism are readily falsifiable. Their rejection or further validation requires measurements of the correlation in cortical responses to odor pairs before and after odor discrimination/association training in two sets of conditions: a control set, in which the response properties of cortical neurons are not altered, and an experimental set, with heightened piriform cortical excitability (i.e. lower activation thresholds). An increase in excitability could be facilitated in a number of ways, for example via stimulation of other brain regions (84) or application of a GABA-A antagonist (85).

**Clinical relevance.** Olfactory deficits are well-known prodromal symptoms of neurological disorders such as Alzheimer’s disease and Parkinson’s disease. In particular, Alzheimer’s disease has been characterized by an increase in olfactory cortex activity, suggesting a possible decrease in the firing threshold of olfactory cortical cells (86, 87). Under such conditions, our results predict that 1) pattern completion would occur at much higher levels than normal for dissimilar odors due to the reversal in the pattern completion trend with initial odor similarity; and 2) pattern separation would be below normal levels due to the weakening in the pattern separation gain. Both these aspects would consequently predict an increase in perceptual similarity of otherwise distinguishable odors as the chief hallmark of olfactory impediment. Generally, Alzheimer’s patients do suffer from loss of odor discrimination ability (88–90); interestingly, in one study, olfactory impairment in patients with Alzheimer’s disease indeed manifested as a perceptual merging of dissimilar odors (91). Our theory then suggests that normal perceptual discrimination ability could be restored by lowering cortical excitability.

**Materials and Methods.** Detailed methods and derivations for the statistical and mechanistic models are provided in Supplementary Information.

**ACKNOWLEDGMENTS.** We thank Jay A. Gottfried for fruitful discussions. DK would also like to thank Minghong Ma and Graeme Lowe for their advice on olfactory bulb anatomy, Eugenio Piasini for use of his server, as well as Max Kelz, Joseph Aicher, Rebecca Li, and Matthew Kerssen for useful discussions and feedback. VB and DK were supported by the Simons Foundation through MMLS grant 400425. GT was supported by the Swartz Foundation.
