

CONTENTS

Figure S1 (related to Figure 1)	<i>Measurement and characterization of odor-evoked behavioral responses</i>
Figure S2 (related to Figure 1)	<i>Behavioral effect of genetically blocking synaptic transmission in Orco-dependent ORNs</i>
Figure S3 (related to Figure 2)	<i>Characterization of the accuracy of the image registration procedure</i>
Figure S4 (related to Figure 2)	<i>Linear correlations between PN and behavioral responses</i>
Figure S5 (related to Figure 5)	<i>PN-PN connectivity inferred from Ca^{2+} imaging data</i>
Figure S6 (related to Figure 7)	<i>Comparison of model predictions for the change in relative valence as a function of the breadth of the olfactory context</i>
Figure S7 (related to Experimental Procedures)	<i>Characterization of the odor delivery system</i>
Table S1 (related to Figure 1)	<i>List of odors used in the study</i>
Table S2 (related to Figure 2)	<i>Mean Ca^{2+} responses of PNs in 37 glomeruli to a set of 84 odors (legend)</i>
Table S3 (related to Experimental Procedures)	<i>List of genotypes used in the study and associated RRIDs</i>
Movie S1 (related to Figure 1)	<i>Reconstruction of a behavioral trial exemplifying aversive and attractive responses. (legend)</i>
Movie S2 (related to Figure 2)	<i>Representative odor response of PNs (legend).</i>
Supplemental Experimental Procedures	
Supplemental References	

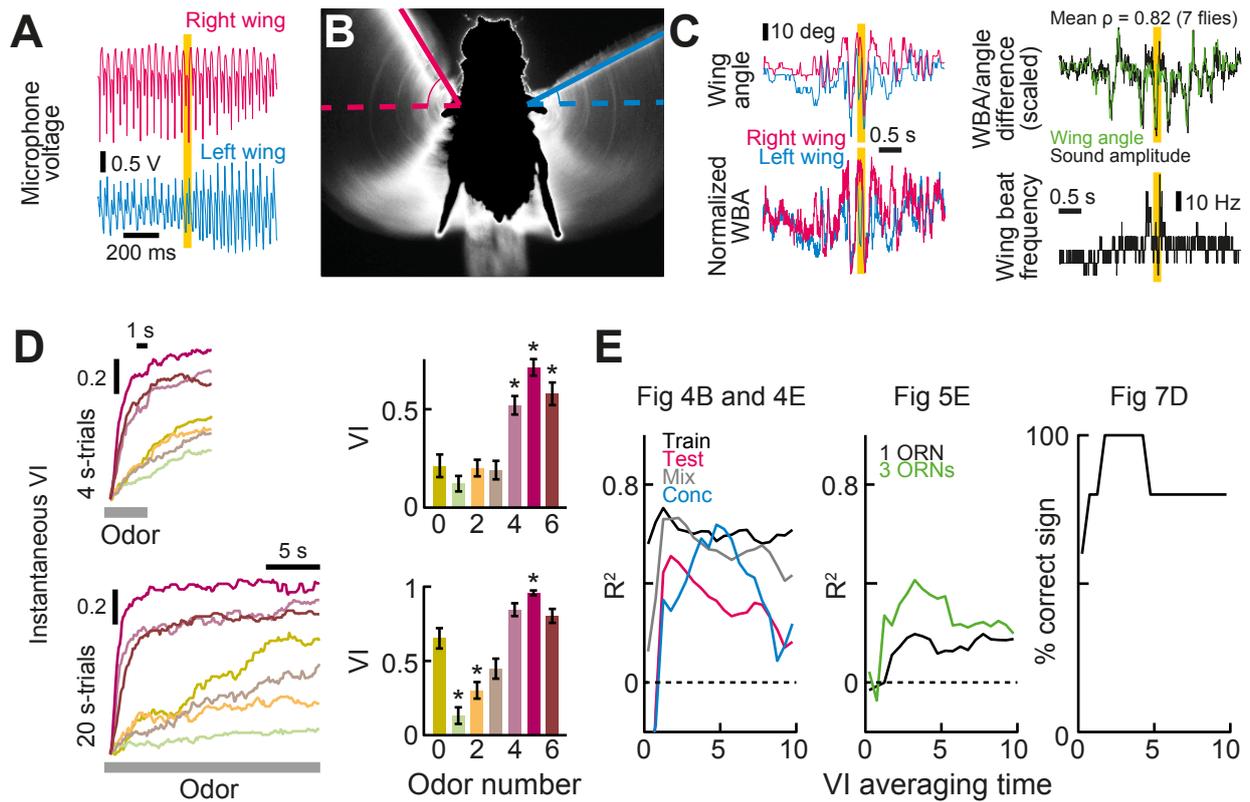


Figure S1 (related to Figure 1). Measurement and characterization of odor-evoked behavioral responses

(A) Sample raw sound data recorded with the left and right microphones. Each wing beat evokes a sharp peak in the microphone voltage with varying amplitude.

(B) Sample snapshot taken using the IR camera and illustration of the left and right wing angles.

(C) Left: wing angles extracted from video (top) and wing-beat amplitudes extracted from sound (bottom) data. Top right: the difference in wing beat amplitudes, which was used to update the heading direction, is highly correlated with the difference in wing angles (average correlation coefficient calculated over 6 s of data in 7 flies, $\rho = 0.82$). Bottom right: the wing-beat frequency is also accessible from sound data. The vertical yellow lines in A and C correspond to the time of the snapshot in B.

(D) Comparison of behavioral responses to 4 s- and 20 s-long odor presentations. Left: instantaneous valence index (VI, proportion of time spent outside of the odor plume) as a function of the time since odor onset. All odors evoke monotonic increases that are faster for aversive (4-6 in right graph) than attractive (1-3) odors. Attractive odors differentiate better from the control odor (0, mineral oil) after extended sampling. Right: the VI is obtained by averaging this quantity in a specified time window (top: 3-4 s, bottom: 15-20 s after odor onset). Error bars denote SEM across flies ($n = 21-22$). Asterisks indicate statistically different values from the mineral oil control (odor 0; Mann-Whitney U-test, Bonferroni corrected, $p < 0.05$). Odor color code is the same as in Figure 1.

(E) Model performance as a function of the time interval used to average VI values. Here, mean VIs are calculated by averaging the instantaneous VI in 0.5 s-time windows, and these values are used to repeat the analyses of Figures 4B and 4E (left), 5E (middle) and 7D (right).

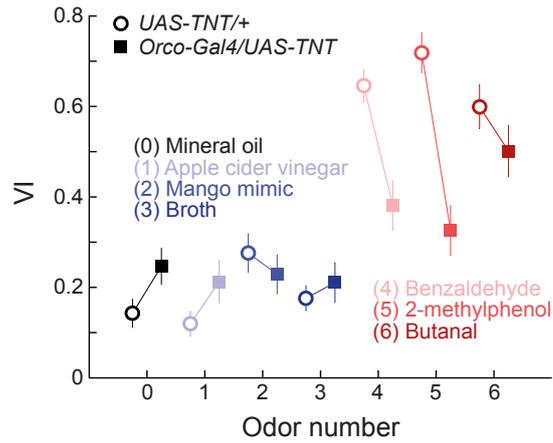
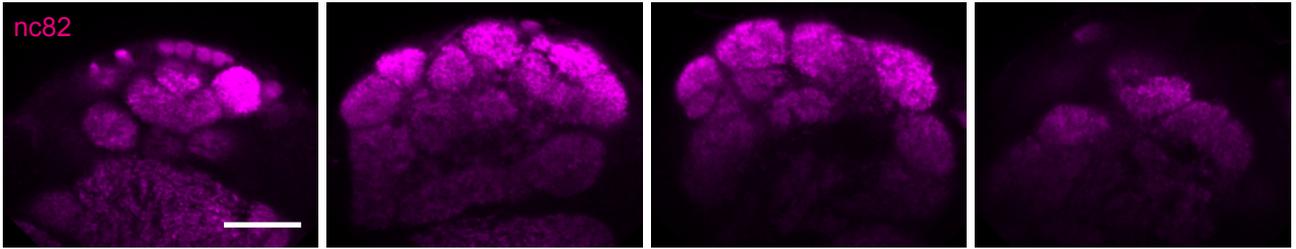


Figure S2 (related to Figure 1). Behavioral effect of genetically blocking synaptic transmission in Orco-dependent ORNs

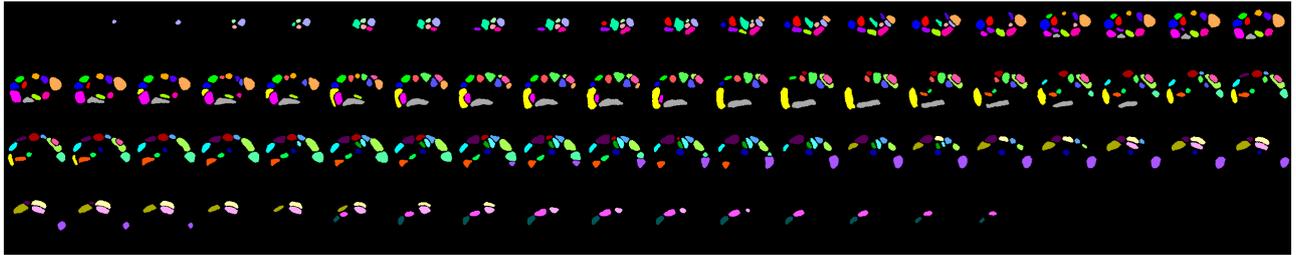
Behavioral responses to mineral oil (black), 3 attractive (blue) and 3 aversive (red) odors are compared across 2 genetic strains: a control strain (*UAS-TNT/+*, open circles) and a strain expressing tetanus toxin (TNT), a blocker of synaptic transmission, in the majority of ORNs (*Orco-Gal4/UAS-TNT*, filled squares). Error bars represent SEM (n = 22 flies for each genotype).

A

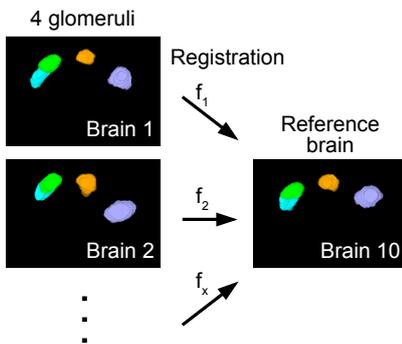
(1) Stain individual brains with antibodies



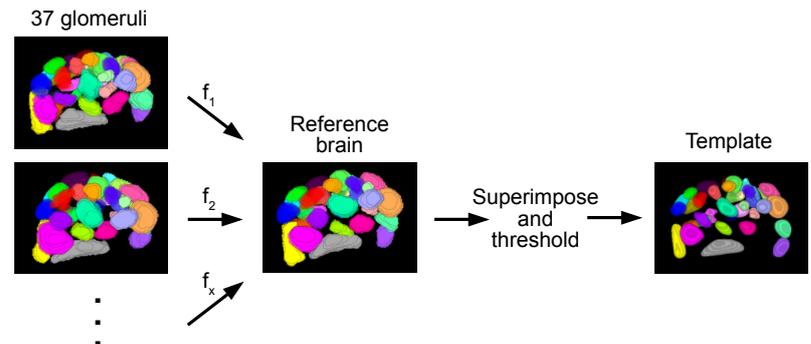
(2) Delineate the 37 NP225-Gal4-positive glomeruli



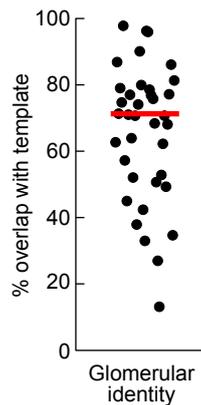
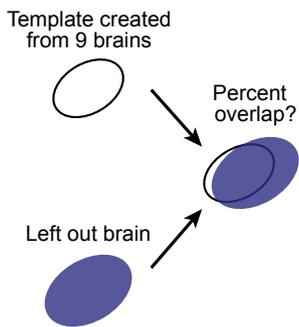
(3) Register 4 guidepost glomeruli



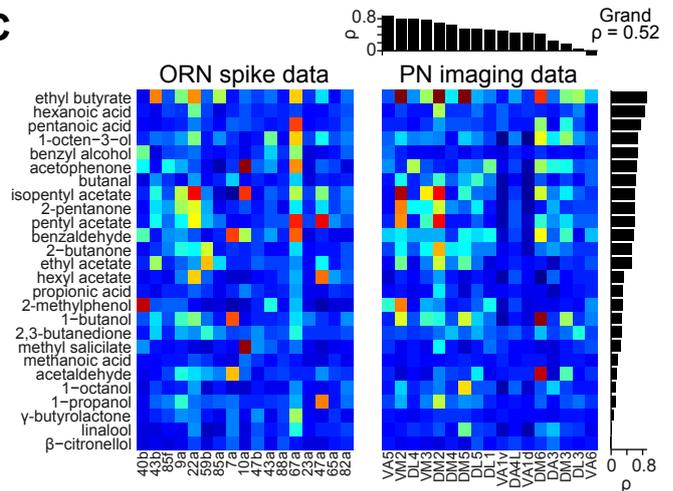
(4) Apply the same transformation to all glomeruli and create the template



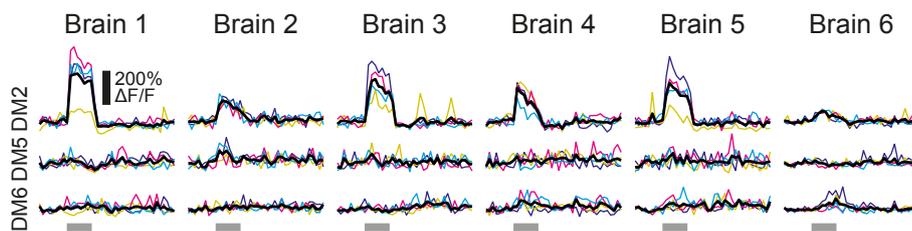
B



C



D



E

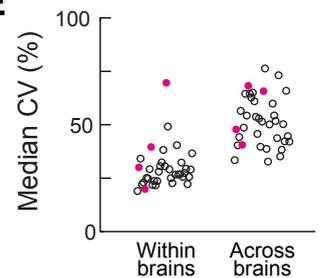


Figure S3 (related to Figure 2). Characterization of the accuracy of the image registration procedure

(A) Creation of the template brain used for registering 2-photon images. (1) Individual brains are stained with antibodies to reveal the morphology of glomeruli (nc82) and transgene expression (anti-GFP, not shown), and imaged at high resolution. The brain was scanned with dorsal side up. Shown are 4 representative slices separated by 18 μm . Scale bar, 20 μm . (2) 37 *NP255-Gal4*-positive glomeruli are manually delineated based on nc82 and anti-GFP signals. (3) Delineated glomeruli in each brain are registered to a reference brain by computing the affine transformation which maximizes the cross-correlation between 4 guidepost glomeruli (DL3, DA2, VM2 and DM5). (4) The 37 delineated glomeruli in all registered brains are superimposed, and thresholded to obtain a template that covers 50% of the mean volume of each glomerulus.

(B) The quality of the registration procedure is assessed by computing, for each glomerulus, the overlap between its volume in individual registered brains and the template. This comparison was repeated 10 times, each time leaving out a different brain. The average overlap over the 10 left-out brains is reported for each glomerulus. The median overlap (red bar), computed across all glomeruli, was 70.8%.

(C) Comparison of PN responses in our data set with the ORN data of Hallem and Carlson (2006), including odors and glomeruli/ORNs that are common to both data sets (arbitrary color scale with warmer colors indicating stronger responses). Bar graphs show the Pearson correlation between both data sets, for each glomerulus and each odor. The Pearson correlation between the entire data sets was 0.52.

(D) Example $\Delta F/F$ traces in response to hexanoic acid in 3 neighboring glomeruli (DM2, DM5 and DM6), for 4 trials of odor application (colored lines), and across all measured brains ($n = 6$). The thick black line denotes the average over 4 trials; the gray bar indicates the period of odor application. The response of DM2 is consistently observed across trials and brains, and no contamination is seen in the other 2 glomeruli, which are not activated by hexanoic acid.

(E) The variability of glomerular responses is summarized using the coefficient of variation (CV). Within-brain variability is assessed by computing, for each glomerulus, the CV of the 4 response trials and taking the median of this quantity across odors and brains. Across-brain variability is assessed by computing trial-averaged responses for each glomerulus, computing their CV across all measured brains and taking the median of this quantity across odors. In both cases, the computation is restricted to odors for which the glomerulus exhibited significant responses. Magenta dots correspond to guidepost glomeruli.

These data suggest that our registration procedure reliably extracts individual glomerular responses without substantial contamination from neighboring glomeruli. We cannot, however, rule out some degree of contamination due to animal-to-animal variability in AL morphology, the extent of which is assessed by the analysis in B. Nevertheless, because morphological variation must occur randomly, it can be safely assumed that no systematic contamination is present in our data. Furthermore, the impact of randomly-occurring contamination can be expected to be minimized after averaging over many brains. To assess the potential impact of such contamination on the result of our analyses more directly, we performed numerical simulations in which contamination was artificially introduced in every glomerulus of every brain by mixing its response with 10% or 20% of the response of one of its 3 nearest neighbors (for each glomerulus, one neighbor was picked at random in each brain). We then repeated our entire analysis and assessed the impact of the contamination on the results. We found that the impact was minimal; in particular, glomerulus weights were always very highly correlated with the original weights of Figure 4.

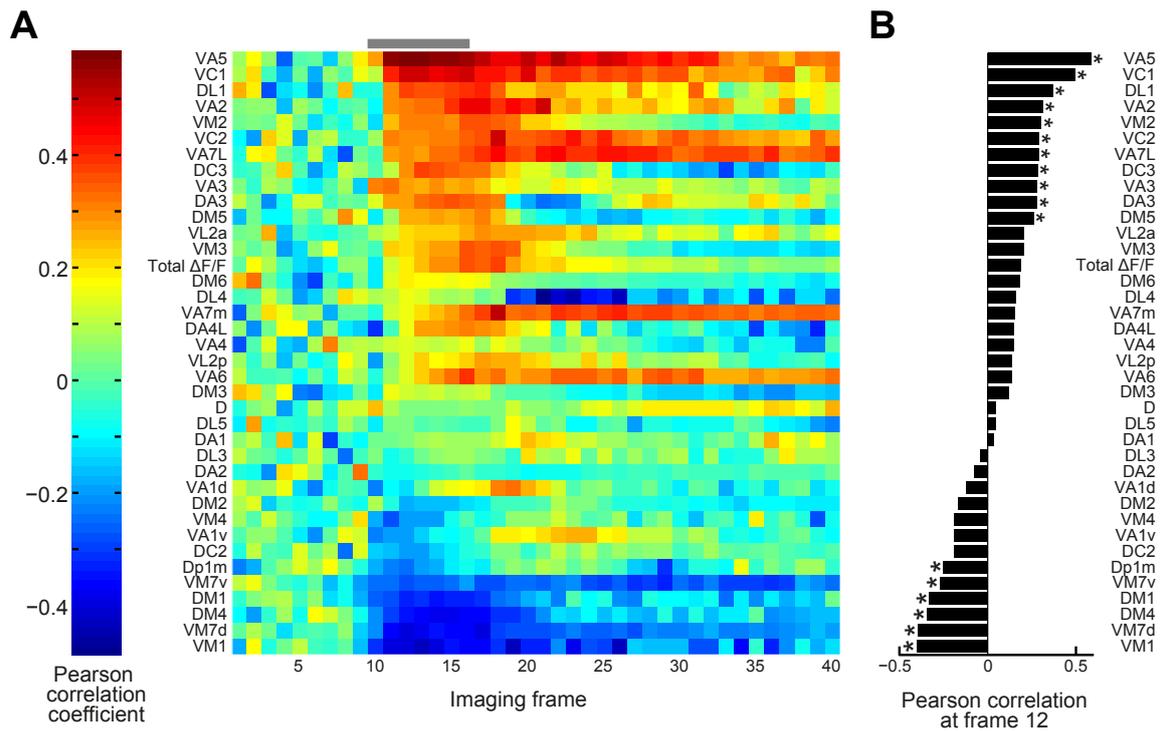


Figure S4 (related to Figure 2). Linear correlations between PN and behavioral responses

(A) Pearson correlation coefficient between PN $\Delta F/F$ activity and VI values, calculated over a data set of pure odors and mixtures (72 odors). Data is color-coded according to the scale bar on the left. Glomeruli are arranged by increasing correlation. The gray bar at the top indicates the period of odor application.

(B) Histogram of the Pearson correlation coefficient at imaging frame 12. Asterisks indicate significant values at the $p < 0.05$ level. Glomeruli are arranged in the same order as in A. The total $\Delta F/F$ (sum of $\Delta F/F$ over all 37 glomeruli) is positively, but weakly correlated with VI, indicating that odor valence is not solely dependent on the overall intensity of PN activation. Note that the rank-order of glomeruli differs from the regression result of Figure 4, as the correlation measures the strength of the linear relationship between a glomerulus' activity and the VI without considering the influence of other glomeruli. In some cases (e.g., VA2, VC1, Dp1m), the apparent sign of the relationship indicated by the correlation may also reverse when other predictors are taken into account.

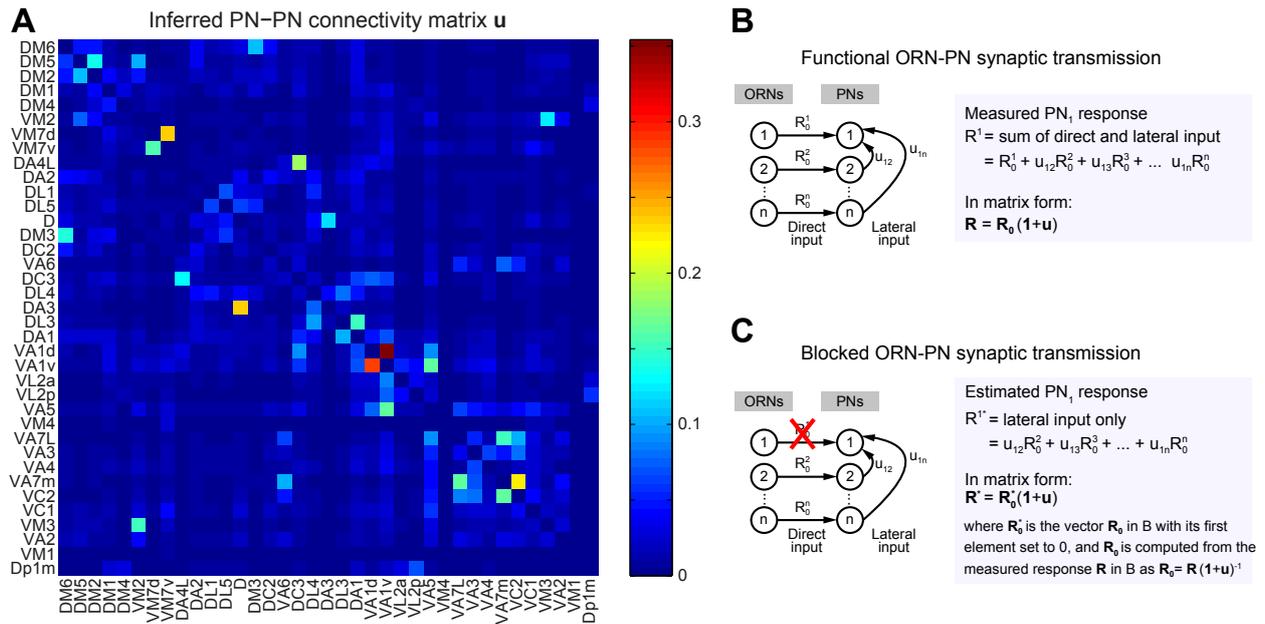


Figure S5 (related to Figure 5). PN-PN connectivity inferred from Ca^{2+} imaging data

(A) Inferred connectivity matrix between 37 glomeruli. The strength of connections is color-coded according to the scale bar on the right.

(B) Model assumed in estimating the connectivity matrix. PNs in each glomerulus receive, in addition to direct ORN input, lateral input from all the other glomeruli, which scales with the strength of their connection. The observed total PN activity is thus the sum of the activity evoked by direct input \mathbf{R}_0 and lateral input $\mathbf{R}_0\mathbf{u}$ where \mathbf{u} denotes the connectivity matrix in A.

(C) Calculation used to estimate PN activity in the case where ORN-PN synaptic transmission is blocked in specific glomeruli. The activity due to direct ORN input, \mathbf{R}_0 , is first calculated from the observed data \mathbf{R} by inverting the relation in B. The elements of \mathbf{R}_0 corresponding to the blocked glomeruli are then manually set to zero, and this is re-convolved with the connectivity matrix to obtain the estimated activity \mathbf{R}^* .

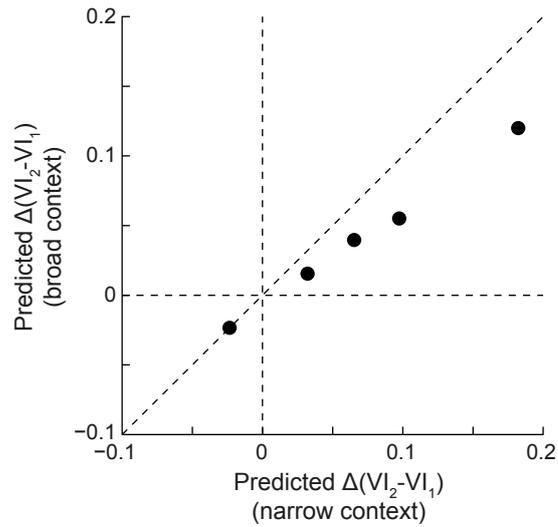


Figure S6 (related to Figure 7). Comparison of model predictions for the change in relative valence as a function of the breadth of the olfactory context

The graph shows theoretical predictions relating to the experimental data for the “same flies” condition of Figure 7E. Flies are presented with a set of 6 odors (context A) in the first half of the experiment, and a different set of 6 odors (context B) in the second half. The relative valence of odor pairs in the first half is computed relative to context A. For the second half, however, it may be computed either relative to context B only (“narrow context”), or, assuming that flies are still affected by the presentation of the previous context, relative to the sum of contexts A and B (“broad” context), leading to different predictions.

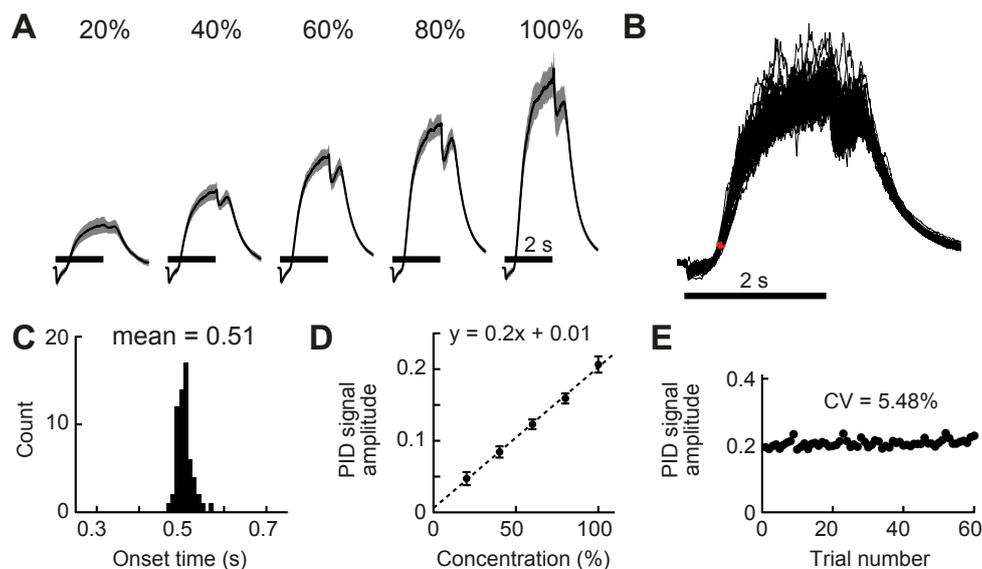


Figure S7 (related to Experimental Procedures). Characterization of the odor delivery system

(A) Average photo-ionization detector (PID) signal time course in response to 4-methyl-cyclohexanol (MCH) mixed with mineral oil at various ratios. Odor is delivered by bubbling the air stream through various combinations of 5 MCH (50% dilution) and 5 mineral oil vials. Labels indicate the ratio of MCH-to-oil vials. Scale bar denotes the period during which solenoid valves were open. Gray shade indicates SEM.

(B) Individual traces corresponding to the 100% condition in A. Overlaid are 60 traces obtained by opening between 1 and 5 MCH-loaded channels, showing that (i) the signal is stable across repetitions, and (ii) the number of open channels does not affect odor delivery significantly. Odor onset times (10% rise time) are shown in red.

(C) Histogram of odor onset time (10% rise time of PID signal, red dots in B) for the 100% condition, showing little variability around its average (0.51 s).

(D) Response amplitude (average PID signal in 1.5-2 s after solenoid valve opening) plotted against the concentration of MCH (ratio of MCH-to-mineral oil vials), showing a linear increase. Dashed line: linear fit. Error bars indicate the standard deviation.

(E) Response amplitude as a function of repeat number for the 100% condition, showing stability across repeats (each MCH-carrying channel was opened between 237 and 276 times across the duration of the experiment).

Pure odors (36 odors)

odor	rank in Fig 1F	solvent	dilution	vendor	odor sets
1-octanol	34	oil	10 ⁻²	Sigma-Aldrich	P ₁ ,T ₂ ,X ₁
3-octanol	32	oil	10 ⁻²	Tokyo chemical industry	P ₂
4-methylcyclohexanol	31	oil	10 ⁻²	Sigma-Aldrich	P ₂
geosmin	19	oil	10 ⁻²	Sigma-Aldrich	P ₄
1-octen-3-ol	13	oil	10 ⁻²	Tokyo chemical industry	P ₅ ,T ₃
1-butanol	17	water	10 ⁻²	nacalai tesque	P ₅ ,T ₂ ,X ₁ ,X ₂
methanoic acid (formic acid)	26	water	10 ⁻²	Sigma-Aldrich	P ₁ ,X ₁ ,X ₂
hexanoic acid	30	water	10 ⁻²	Sigma-Aldrich	P ₁
pentanoic acid (Valeric acid)	28	water	10 ⁻²	Sigma-Aldrich	P ₂
2-oxopentanoic acid	14	water	10 ⁻²	Sigma-Aldrich	P ₃ ,X ₁
propionic acid	7	water	10 ⁻²	Wako	P ₄
ethyl butyrate	20	oil	10 ⁻²	Sigma-Aldrich	P ₂ ,T ₂ ,T ₃
hexyl acetate	33	oil	10 ⁻²	Wako	P ₃
isopentyl acetate	12	oil	10 ⁻²	Wako	P ₃ ,T ₂ ,X ₂
ethyl acetate	11	oil	10 ⁻²	Sigma-Aldrich	P ₄ ,T ₂ ,X ₂
pentyl acetate	4	oil	10 ⁻²	Sigma-Aldrich	P ₅
benzaldehyde	37	oil	10 ⁻²	Sigma-Aldrich	P ₂ ,T ₁
2-methylphenol (o-cresol)	38	oil	10 ⁻²	Sigma-Aldrich	P ₂ ,T ₁
acetophenone	35	oil	10 ⁻²	Sigma-Aldrich	P ₁ ,X ₁
benzyl alcohol	15	water	10 ⁻²	Wako	P ₄ ,X ₁
methyl salicylate	8	oil	10 ⁻²	Sigma-Aldrich	P ₅
2,3-butanedione (Diacetyl)	9	water	10 ⁻²	nacalai tesque	P ₂
2-pentanone	1	oil	10 ⁻²	Wako	P ₃ ,T ₃
2-butanone	3	water	10 ⁻²	Sigma-Aldrich	P ₅ ,T ₃ ,X ₂
butanal (butyraldehyde)	36	oil	10 ⁻²	nacalai tesque	P ₂ ,T ₁
acetaldehyde	29	water	10 ⁻²	Wako	P ₃
β-citronellol	16	oil	10 ⁻²	Wako	P ₄
linalool	23	oil	10 ⁻²	nacalai tesque	P ₄ ,T ₂
γ-butyrolactone	24	water	10 ⁻²	Sigma-Aldrich	P ₁
3-methylthio-1-propanol	27	water	10 ⁻²	Sigma-Aldrich	P ₂ ,X ₂
phenylethylamine	25	water	10 ⁻²	Wako	P ₃
apple cider vinegar	5	NA	1	Mizkan	P ₂ ,T ₁
mango mimic ^a	18	water	10 ⁻²	see Other components	P ₂ ,T ₁ ,T ₃
broth	6	water	10 ⁻²	Yamasa	P ₂ ,T ₁
vinegar mimic ^b	22	water	10 ⁻²	see Other components	P ₁ ,T ₃
banana essence	2	water	10 ⁻²	Narizuka corporation	P ₅

Mixtures (3 x 3 = 9 mixture sets, 36 different mixtures excluding pure components)

odor A	odor sets	odor B	odor sets
apple cider vinegar	M ₁ ,M ₂ ,M ₃	benzaldehyde	M ₁ ,M ₄ ,M ₇
mango mimic	M ₄ ,M ₅ ,M ₆	2-methylphenol	M ₂ ,M ₅ ,M ₈
broth	M ₇ ,M ₈ ,M ₉	butanal	M ₃ ,M ₆ ,M ₉

odor A:odor B = 100:0, 80:20, 60:40, 40:60, 20:80, 0:100

Concentration series (12 odors, excluding the concentration used in the pure odor set)

odor	solvent	dilution	steps	odor sets
apple cider vinegar	water	10 ⁻¹ ~ 10 ⁻³	10 ⁻¹	C ₁
mango	water	10 ⁻³ ~ 10 ⁻⁵	10 ⁻¹	C ₂
benzaldehyde	oil	10 ⁻³ ~ 10 ⁻⁵	10 ⁻¹	C ₁
2-methylphenol	oil	10 ⁻³ ~ 10 ⁻⁵	10 ⁻¹	C ₂

Other components (for mango mimic and vinegar mimic)

odor	vendor
ethanol	Junsei chemical
acetic acid	Wako
2-phenylethanol	Wako
ethyl acetate	Sigma-Aldrich
acetoin	Sigma-Aldrich

- a. Ethanol, acetic acid, and 2-phenyl ethanol were mixed at a ratio of 1:22:5 (volume) as described in Zhu et al. (2003).
b. We modified the composition described in Becher et al. (2010) by taking vapor pressure and density into account, and dissolved acetic acid, 2-phenyl ethanol, acetoin, and ethyl acetate in ethanol at a ratio of 1:13.48:0.21:0.012 (volume).

Table S1 (related to Figure 1). List of odors used in the study

The table lists the 84 odors used in the study. Odor sets are divided into 5 categories: pure odors (P), mixtures (M), concentration (C), TNT (T) and context (X).

Table S2 (related to Figure 2). Mean Ca²⁺ responses of PNs in 37 glomeruli to a set of 84 odors

The table gives the mean PN activity ($\Delta F/F$ averaged over odor application period, frames 10-17) in 37 glomeruli in response to the 84 odors used in the study, including the 36 odors displayed in Figure 2.

Figure	Genotype
1,2B-2E,3,7,S1,S3,S4,S6	<i>NP225-Gal4,UAS-IVS-GCaMP6f(attP40)/UAS-IVS-GCaMP6f(attP40)</i>
2A	<i>NP225-Gal4/UAS-mCD8::GFP</i>
5A,5D-5G	Blockade of single glomerulus:
	F1 of <i>Or-Gal4 x UAS-TNT-E2</i>
	<i>Or33c-Gal4</i>
	<i>Or42b-Gal4</i>
	<i>Or59b-Gal4</i>
	<i>Or83c-Gal4</i>
	<i>Or85a-Gal4</i>
	<i>Or85d-Gal4</i>
	<i>Or92a-Gal4</i>
	<i>UAS-TNT-E2/+</i> (control)
	Blockade of three glomeruli:
	<i>Or85a-Gal4/UAS-TNT-E2;Or33c-Gal4/Or83c-Gal4</i>
	<i>Or85d-Gal4/UAS-TNT-E2;Or59b-Gal4/Or92a-Gal4</i>
5B,5C	<i>Or42b-Gal4/UAS-TNT-E2;GH146-QF[#53],QUAS-GCaMP3/QUAS-GCaMP3</i>
	<i>Or85a-Gal4/UAS-TNT-E2;GH146-QF[#53],QUAS-GCaMP3/GH146-QF[#53]</i>
6	<i>UAS-IVS-CsChrimson.mVenus(attP40);Or59b-Gal4</i>
S2	<i>Orco-Gal4/UAS-TNT-E2</i>
	<i>UAS-TNT-E2/+</i> (control)

Genotype	RRID
<i>NP225-Gal4</i>	DGGR_112095
<i>Or33c-Gal4</i>	BDSC_9966
<i>Or42b-Gal4</i>	BDSC_9971
<i>Or59b-Gal4</i>	BDSC_22898
<i>Or83c-Gal4</i>	BDSC_23132
<i>Or85a-Gal4</i>	BDSC_23133
<i>Or85d-Gal4</i>	BDSC_24148
<i>Or92a-Gal4</i>	BDSC_23140
<i>Orco-Gal4</i>	BDSC_26818
<i>UAS-IVS-GCaMP6f(attP40)</i>	BDSC_42747
<i>UAS-IVS-mCD8::GFP(attP40)</i>	BDSC_32186
<i>UAS-TNT-E2</i>	BDSC_28837
<i>UAS-IVS-CsChrimson.mVenus(attP40)</i>	BDSC_55135
<i>GH146-QF[#53]</i>	BDSC_30015
<i>QUAS-GCaMP3</i>	BDSC_52231

Table S3 (related to Experimental Procedures). List of genotypes used in the study and associated RRIDs

All the flies are from the Bloomington Stock Center except for the following stocks: Michael Dickinson Laboratory wild type and *w¹¹¹⁸* are gifts from Rachel Wilson; *NP225-Gal4* is a gift from Hiromu Tanimoto; *GH146-QF[#53]* is a gift from Christopher Potter.

Movie S1 (related to Figure 1). Reconstruction of a behavioral trial exemplifying aversive and attractive responses

The movie shows a sequence of snapshots taken using the infrared camera along with a reconstruction of the visual stimulus delivered to the fly during the trial. An aversive odor (benzaldehyde) is delivered at $t = 6.5$ s, and an attractive odor (apple cider vinegar) at $t = 16.5$ s.

Movie S2 (related to Figure 2). Representative odor response of PNs

Representative response to 3-octanol in *NP225-Gal4* positive PNs expressing GCaMP6f imaged with two-photon microscopy. Thirty different focal planes aligned from top-left to bottom-right along the dorsoventral axis are separated by $3 \mu\text{m}$. $\Delta F/F$ signal (pseudo-color, average of four trials) is overlaid on the structure of the AL (grayscale). A white bar indicates the 4 s odor application period and measures $10 \mu\text{m}$. The movie is sped up 6 times.

SUPPLEMENTAL EXPERIMENTAL PROCEDURES

Fly Stocks

Flies were raised on conventional cornmeal agar medium under a 12 hr light/ 12 hr dark cycle at 25 °C. The density of culture was controlled to be sparse. All experiments were performed on adult female flies, 2-4 days after eclosion. Therefore, our results technically do not apply to males. Flies were starved for 4-6 h with water prior to experiments. The following stocks were used: Michael Dickinson Laboratory wild type (Bhandawat et al., 2010), *UAS-GCaMP6f(attP40)* (Chen et al., 2013), *UAS-TNT-E2* (Sweeney et al., 1995), *UAS-mCD8::GFP*, *NP225-Gal4* (Tanaka et al., 2012; Thum et al., 2007), *GH146-QF* (Potter et al., 2010), *QUAS-GCaMP3* (Tian et al., 2009), *UAS-Chrimson* (Klapoetke et al., 2014), and *Or-Gal4* lines (Couto et al., 2005; Fishilevich and Vosshall, 2005; Larsson et al., 2004) including *Or33c-Gal4*, *Or42b-Gal4*, *Or59b-Gal4*, *Or83b-Gal4* (*Orco-Gal4*), *Or83c-Gal4*, *Or85a-Gal4*, *Or85d-Gal4*, *Or92a-Gal4*. For Chrimson experiments, 200 mM all-transretinal dissolved in ethanol was mixed with fly food and fed to the flies for 3 days prior to experiments. All stocks were back-crossed 6 generations to Dickinson Laboratory wild type flies carrying w^{1118} except for Q system flies. Detailed genotypes of flies used in each experiment are listed in Table S3.

Olfactory Stimulation

Odors were delivered by passing an air stream through 4 ml of odorant solution (using either mineral oil or water as a solvent) in a glass vial. The air stream (250 ml/min) was split into 16 parallel channels each lined up with a single odor vial and a solenoid valve (100E1-SR, Koganei Corporation) that regulated the open/closed state of the channel. The outputs of all channels were pooled, mixed into the main air stream (1.55 l/min), and directed into the behavioral chamber using $\varnothing 6$ mm Teflon tubing (main tube). A small portion of this air stream was diverted from the main tube through a $\varnothing 2$ mm outlet tube and delivered frontally to the fly. This configuration reduced the impact of transient changes in air pressure caused by the switching of solenoid valves, and was critical not to disturb flight at the onset of odor delivery. The tip of the delivery outlet was placed 10 mm away from the fly. The remainder of the air stream, which was not delivered to the fly, was directed out of the chamber and cleared using a fume hood fan unit (3-4064-11, AS ONE). Because the quantity of air reaching the fly decreased with increasing suction of the fan unit, suction power was calibrated to obtain an air flow of 0.3 m/s at the fly location, using an air flow sensor (QB-5, Tohnic). Odors delivered inside the chamber were actively removed by a suction tube placed behind the fly. Binary mixtures of two odors A and B were obtained by simultaneously circulating air through 5 channels, each loaded with either A or B, and varying the relative number of A and B channels. This method was chosen because the concentration of delivered odor, as measured using a photo-ionization detector (200B miniPID, Aurora Scientific Inc.), varied linearly (Figure S7). To avoid odor contamination, Teflon was used for all parts that came in contact with odorized air, and tubing was periodically washed with alcohol, rinsed with purified water, and dried with clean air. A total of 84 odors were used in the study (Table S1).

Behavioral Experiments

Flies were cold-anesthetized on a Peltier device and tethered to a stainless steel pin (Austerlitz minutiens $\varnothing 0.1$ mm) between the head and thorax, using ultraviolet-curing adhesive (NOA 63, Norland). Tethered flies were placed in a glass vial equipped with a piece of filter paper soaked with purified water to prevent dehydration, starved, and transferred to the flight-simulator arena. The setup consisted of an odor delivery apparatus and a 24x56 array of green LEDs (Reiser and Dickinson, 2008) arranged in a half circle (Mettrix Technology Corp.). The display spanned ~ 210 deg horizontally, and ~ 56 deg vertically below the fly. The entire setup was enclosed in an opaque container to prevent light entry, and the tethered fly was illuminated with infrared LEDs to allow visualization using a camera (Lu070M, Lumenera corporation).

Flight behavior was monitored using two microphones (AT9904 electret condenser microphones, audio-technica) positioned laterally ~ 1 mm from the tip of the extended wing on either side of the fly, whose outputs were amplified (AT-MA2 amplifier, audio-technica), digitized (NI 9215, National Instruments) and analyzed in real-time using a desktop computer (Optiplex 980, Dell) to extract the turning direction and speed, which were used to update the visual and olfactory stimuli in closed-loop. The visual stimulus was essential to increase flight reliability, and consisted of vertical gratings with a spatial frequency of 60 deg^{-1} . The panel of 84 olfactory stimuli was divided in sets of 6 odors (Table S1) and a single set was applied in each experiment (except for the data in Figure 7E, “same flies” condition, where two sets were applied consecutively). The protocol consisted of 15 blocks, in which each of the 6 odors, as well as a control odor (mineral oil), were applied in randomized order, up to a fixed duration (4 s unless otherwise stated) and in a restricted spatial region (45 deg centered at the fly location at the time of odor contact). Odor application was terminated when the fly exited this spatial region, but was re-initiated if the fly re-entered the region within the application period. For optogenetic activation experiments, a 637 nm laser (HL63142DG, Thorlabs) was positioned in the setup and controlled in closed-loop in the same way as odorant stimuli. Before each experiment, the laser was adjusted to target the antennal region. Inter-trial-interval was 10 s unless otherwise stated. When no odor was applied (including inter-trial intervals), air was presented in order to maintain a fixed air flow. The random sequence of odor presentation was fixed across experiments and chosen not to present the same odor consecutively at the transition between successive blocks. After post-hoc data analysis, data from individual flies were used for analysis if the flying percentage during the entire experiment exceeded 50% (resulting in a rejection rate of $\sim 1/10$). For each data set, experiments were conducted until data were collected from 20 flies. This number was chosen on the basis of preliminary experiments to obtain a sufficient signal-to-noise ratio.

Closed-loop control of the visual display and solenoid valves was achieved using a data-acquisition board and modules (NI cDAQ9178, NI 9215, NI 9264, National Instruments) and custom software written in Matlab (Mathworks), Java and C. Briefly, microphone signals were acquired continuously in chunks of 5 ms, and signal amplitude in each chunk was computed as the difference between the maximum and the minimum values. These amplitudes were filtered by calculating the median over the 3 most recent values and (i) summed to obtain the flight strength s , and (ii) standardized (using mean and standard deviation values computed over the previous block of trials; for the first block, running estimates computed before the experiment were used) to obtain the standard left and right wing-beat amplitudes w_L and w_R . Turns during flight were assessed by monitoring the difference in wing-beat amplitudes, a proxy for yaw torque (Tammero et al., 2004). An increment $\Delta\theta$ in angular position was registered if the difference in standard wing-beat amplitudes $\Delta w = w_R - w_L$, after multiplication by the flying strength, exceeded the value of its standard deviation $\sigma_{\Delta w}$ (computed over the previous block of trials), above which the increment was proportional to Δw , i.e., $\Delta\theta = \gamma \text{sign}(\Delta w) \max(0, |\Delta w|s - \sigma_{\Delta w})$, with a coupling coefficient $\gamma = 0.375$ to yield units of degrees. Angular position was computed by cumulatively summing these increments, and odorant and visual stimuli were updated every 5 ms in accordance with the current position. The validity of computing the turning propensity from microphone data was verified by comparing with the wing angles obtained from video data; the two methods yielded highly similar results (Figure S1).

A total of 84 odors were tested with the set repartition listed in Table S1. The analysis in Figures 4B-4D was conducted using 5 pure odor sets P_1 - P_5 and 9 mixture sets M_1 - M_9 . The final model was thus estimated using these 14 sets but none of the other data. For Figure 4E, two additional test sets, concentration sets C_1 - C_2 , were used. For these sets, odors evoking strong behavioral responses were chosen because of the expectation that lowering the concentration would reduce the amplitude of the responses. For Figure 5, single-ORN experiments were all conducted using the same odor set, T_1 , with the exception of DM5 which was additionally tested with T_2 . Triple-ORN experiments were conducted using sets T_2 (DM5+DC3+VC1) and T_3 (DM4+VA2+VA4). For Figure 7D, additional experiments were conducted using sets X_1 and X_2 , and results were pooled with previous experiments using sets P_1 and P_5 to compare relative valences. For Figure 7E, “same flies” condition, experiments were

conducted by presenting sequentially X_1 and X_2 , X_1 and P_1 , or X_2 and P_5 , with the exact same protocol as in other experiments except that odors were changed after the 8th block.

Two-photon Imaging

Individual starved flies were attached to a custom recording plate modified from Maimon et al. (2010) with ultraviolet-curing adhesive (NOA 63, Norland) while cold-anesthetized. Flies were attached at their anterodorsal end of the thorax, posterodorsal end of the head, and dorsal edges of the compound eyes. A small hole was made on the plate to access the head cuticle for dissection, with most of the fly body remaining below the plate. The plate was designed using computer-aided design software (Rhinoceros 5.0, Robert McNeel and Associates) and milled out of Delrin plastic with an MDX-540SA milling machine (Roland DG) controlled by computer-aided manufacturing software (Craft MILL, C & G Systems). After attaching the flies, saline bubbled with 95% O_2 /5% CO_2 was added on the plate, and the head cuticle was removed with forceps to expose the brain of *NP225-Gal4,UAS-GCaMP6f/UAS-GCaMP6f* flies. Saline contained (in mM): 103 NaCl, 3 KCl, 5 N-tris(hydroxymethyl) methyl-2-aminoethane-sulfonic acid, 8 trehalose, 10 glucose, 26 $NaHCO_3$, 1 NaH_2PO_4 , 1.5 $CaCl_2$, and 4 $MgCl_2$ (osmolarity adjusted to 270-275 mOsm). Saline was perfused at a rate of 2 ml/min during the recording. The muscle of the frontal pulsatile organ (number 16) was severed to alleviate brain movement.

The entire right AL was imaged with a two-photon laser scanning microscope (LSM 710 NLO, Zeiss) equipped with a water immersion objective lens (W Plan-Apochromat, 20x, numerical aperture 1.0). The fluorophore was excited with a titanium:sapphire pulsed laser (Chameleon Vision II, Coherent) mode-locked at 930 nm. Laser power was set between 3-6 mW at the back aperture of the objective lens. Fluorescence was collected with an internal GaAsP detector through a bandpass emission filter (BP470-550). A piezo positioner (P-725.2CD PIFOC, PI) moved the objective lens vertically, allowing the system to cover 33 optical slices separated by 3 μm in 595 ms (18 ms/slice). Such three-dimensional (3D) images of the AL were acquired repetitively during functional imaging, corresponding to volume scanning at a rate of ~ 1.7 Hz. After applying digital zoom, the pixel size in a slice was 1.384 x 1.384 μm , which was sufficient to resolve individual glomeruli measuring on average ~ 10 μm in diameter.

The odor delivery system was identical to that used in behavioral experiments, and the properties of both systems were scrupulously matched. In each experiment, an odor set consisting of 15 pure odors or 6 odor mixtures was applied. Each odor in the set was presented for 4 s, with an inter-trial-interval of 1 min between odors. Odor valves opened in frame 9 and closed in frame 16. Each set was presented 4 times, and the order of odor presentation was randomized in each block. The sequence of odor presentation was fixed across experiments of the same type and chosen not to present the same odor consecutively at the transition between two successive blocks. After collecting odor responses, 1 μM tetrodotoxin was added to the saline and a high-resolution 3D image of the AL comprising 99 optical slices (1 μm interval, 0.346 x 0.346 μm pixel size) was acquired for off-line identification of guide-post glomeruli (see Data Analysis section).

Immunohistochemistry and Structural Imaging

Brains were dissected out from the head capsule in phosphate-buffered saline (PBS) and fixed with 4% paraformaldehyde in PBS for 90 min on ice. Each brain was washed 3 times over an hour and incubated in blocking solution containing 5% normal goat serum in PBST (0.2% Triton-X in PBS) for 30 min. Primary antibodies were added and incubated at 4 $^{\circ}C$ for ~ 48 hr. After removal of antibodies and washing over an hour, the brain was incubated in a solution of secondary antibodies at 4 $^{\circ}C$ for 24 hr. The brain was then washed, immersed in PBS, and positioned with its dorsal side up on a MAS-coated glass slide (Matsunami). Images were collected with an LSM 710 NLO confocal microscope (Zeiss) using a water immersion lens. The only exception was the brain in Figure 2A, which was immersed in Vectashield (Vector laboratories), positioned anterior side up on a glass slide, sealed with a cover glass, and imaged with a FV1000 confocal microscope (Olympus) using an oil immersion lens. The following

antibodies were used at the indicated dilution: mouse nc82 (1:20, Developmental Studies Hybridoma Bank at Univ. of Iowa, RRID: AB_528108), rat anti-GFP (1:1000, nacalai tesque, RRID: AB_10013361), anti-mouse CF633 (1:250, Biotium, RRID: AB_10556971), and anti-rat CF488 (1:250, Biotium, RRID: AB_10557403).

ORN Recordings

ORN recordings were performed as previously described (Kazama and Wilson, 2009). Ab2 sensilla were identified on the basis of morphology, number of distinct spikes and response to ethyl acetate.

Data Analysis

Data were analyzed with custom codes written in Matlab (Mathworks) unless otherwise stated.

Behavioral data analysis. Valence indices (VIs) for individual flies were obtained by aligning flight trajectories to the time of odor contact (0.5 s after odor valve opening, see Figure S7) and calculating the proportion of time spent outside the odor plume for each odor. The instantaneous VI was calculated in every 5 ms time bin, and the mean VI (simply referred to as VI throughout the text) was calculated as an average over the last 1 s of odor presentation period. The mean and standard error were then computed by pooling all tested flies. Δ WBA in Figure 1G was calculated using the absolute difference in the standardized wing-beat amplitudes (c.f. Behavioral Experiments section) and averaged in 100 ms bins before pooling over tested flies. Trials in which flight was interrupted were excluded from the analysis.

Two-photon imaging data analysis. To quantify PN odor responses in individual glomeruli, images were analyzed in three steps: correction for brain motion, creation of template glomeruli, and extraction of fluorescence changes in each glomerulus using the template.

To correct for brain motion, each 3D image resulting from a full AL scan was re-aligned to the high-resolution 3D image with sub-pixel precision. Re-alignment was achieved by shifting the target image in a Cartesian coordinate system so as to maximize the cross-correlation between the two images.

To extract fluorescence changes in individual glomeruli in an efficient and objective manner, we constructed a template AL against which all images were registered. To obtain the template, 37 *NP225-Gal4*-positive glomeruli were delineated in immunostained brains based on nc82 and anti-GFP signals using the Fiji (Schindelin et al., 2012) segmentation editor (Figure S3). Delineated glomeruli in different brains were registered to the reference brain by optimizing an affine transformation to maximize the cross-correlation between four guidepost glomeruli (DL3, DA2, VM2, and DM5), and applying this transformation to the whole brain (Figure S3). These particular glomeruli were selected as guideposts because they are strongly labeled by *NP225-Gal4*, jointly cover a wide surface of the AL, and because four is the minimal number of points required to determine a rigid body. A larger number of guidepost glomeruli were not used because they would collectively resemble a sphere, in which case the cross-correlation between the two images may exhibit several maxima. Registered glomeruli were superimposed and thresholded to obtain the template AL, in such a way that each glomerulus in the template covers 50% of the mean volume of the corresponding glomerulus. This moderate threshold volume was chosen to increase the probability that template glomeruli fall within the boundaries of actual glomeruli. Using this procedure, the median overlap computed across glomeruli was 70.8% with respect to the template volume (Figure S3).

At the end of each experiment, images were registered to the template AL by first delineating the four guidepost glomeruli using the high-resolution image (see Two-photon imaging section). The affine transformation maximizing the cross-correlation between the four glomeruli and the template was then computed and applied to all images. Data from individual brains was discarded without further analysis if the correlation with the template was lower than 0.6 in the 4 guidepost glomeruli. Finally, the template volume was used to calculate the relative change in fluorescence $\Delta F/F = (F_t - F_{\text{baseline}})/F_{\text{baseline}}$ in each glomerulus. Table S2 summarizes the mean response of each of the 37 glomeruli to the 84 odors in our data set. For the classification analysis of Figure 2E, an error-correcting output

codes (ECOC) multiclass model was trained using the `fitcecoc.m` function in Matlab. At each time frame, a classifier was trained using the responses of 37 glomeruli to 84 odors in 3 of the 4 imaging trials, and the data from the remaining trial was used to test classification performance.

Decoding analysis. Raw PN $\Delta F/F$ data were pre-processed as follows. For each imaging trial, a Mann-Whitney U-test ($p < 0.01$) was used to determine the glomeruli whose response (frames 12-15) differed significantly from the baseline (frames 2-8). Glomeruli that responded significantly and with the same mode (excitation or inhibition) in more than half of the trials were marked as responding, and their mean response was computed by averaging over the corresponding trials. For other glomeruli, the mean response was computed by averaging over non-significant trials. This procedure was employed because some glomeruli showed occasional failures in their responses; however, including all imaging trials had no qualitative impact on the result of our analyses.

For the normalization model, PN data was standardized in each glomerulus using mean and standard deviation values computed independently for each individual set of 6 odors, by pooling trial-averaged $\Delta F/F$ responses (imaging frames 12 to 16) over all 6 odors. Importantly, the statistics of glomerular activity are local quantities determined by the history of olfactory stimuli: because a different group of 6 odors was tested in each set of experiments, the same normalization parameters were used for odors in the same group, whereas different parameters were used for odors in different groups. VI values were similarly centered by subtracting their mean value over the 6 odors.

We fit the model using linear regression. Because of the large number of predictors, it was necessary to use a regularization procedure to avoid overfitting the training data. We used partial least-squares (PLS) regression, a technique known to be efficient in problems with many predictors and robust to collinearity in the predictors (Abdi, 2010). A number of regularization methods, including methods that favor sparse readout weights (e.g., LASSO), were also tested but consistently yielded poorer results. Briefly, the PLS method linearly transforms the set of predictor variables into a set of explanatory factors and regresses the dependent variable on the factors. The factors are calculated similarly as in principal component analysis, except that whereas principal component analysis finds directions of maximum variance in the predictors, PLS searches for directions of maximal covariance between the predictors and the dependent variable. Regularization is achieved by limiting the number of factors included in the regression, which can be achieved by cross-validating using a validation set. In our analysis, we used the following cross-validation procedure. In each validation round, one of the 9 binary mixture sets was taken as the test set, and was therefore excluded from the fitting procedure. The 4 other mixture sets that shared a common odor with the test set were also excluded from the fitting procedure to avoid training/test set contamination. The remaining 4 mixture sets, which did not share a common odor with the test set, were used for validation. The training set comprised the 30 pure odors that did not appear in the mixtures. We computed the PLS factors by pooling the training and validation sets using the `plsregress` function in Matlab. We then regressed the VI values on the factors using the training data only, by means of the `fitlm` function, and evaluated the performance of the resulting model on the validation set by calculating the R^2 coefficient. The optimal number of factors was chosen to maximize the validation R^2 . This resulted in 2-8 retained factors (average 3.2), indicating that a significant amount of regularization was necessary.

After each validation round, predicted VI values for the test mixture (6 values) were computed using the fitted model. At the end of the 9 validation rounds, predicted values for the 9 mixture sets were pooled together (54 values), and prediction performance was assessed using the prediction R^2 , defined as $R^2 = 1 - E_{\text{pred}}^2 / \sigma^2$, where E_{pred}^2 is the mean-square prediction error calculated over the 9 mixture sets, and σ^2 is the variance of the test data. This coefficient was chosen because it provides a measure of the prediction error relative to the typical variation in the data, and is thus a quantitative measure of prediction performance that is directly comparable across multiple test sets. It should be noted that, unlike the conventional coefficient of determination (R^2 of fit), the prediction R^2 may

take negative values if the prediction error is larger than the typical variation in the data, which may happen, for example, if the predictions are shifted from the data by a constant offset.

The statistical significance of prediction R^2 values was assessed using a permutation test. 500 surrogate PN data sets were generated by randomly shuffling glomerular labels, and the entire decoding analysis was repeated for each set. Upper bounds on the p-values for the reported statistics were estimated from the distribution obtained with the surrogate data.

For the data of Figure 4C, the best subsets of each size were defined as those that performed best on the four data sets (training, validation, test and concentration) simultaneously. In practice, we selected the subset that maximized the quantity $\sum_i (r_i)^2 - \sum_{i>j} (r_i - r_j)^2$, where r_i denotes the R^2 coefficient for data set i , and indices i and j run over the 4 data sets.

For the analysis of TNT data, PN-PN connections were inferred from the noise correlation structure of Ca^{2+} signals in the period of baseline activity preceding odor application (imaging frames 1 to 8, Figure S5). PN responses \mathbf{R} were modeled using the equation $\mathbf{R} = \mathbf{R}_0(\mathbf{1} + \mathbf{u})$ where \mathbf{R}_0 represents the portion of PN activity due to direct ORN drive, $\mathbf{1}$ is the identity matrix, and \mathbf{u} is the PN-PN connectivity matrix. Given that ORNs fire independently in the absence of odorant stimulus (Kazama and Wilson, 2009), the connectivity matrix \mathbf{u} was calculated by jointly diagonalizing the covariance matrices of PN activities in the considered imaging frames, using a gradient descent algorithm.

SUPPLEMENTAL REFERENCES

- Abdi, H. (2010). Partial least squares regression and projection on latent structure regression (pls regression). *WIREs Comp Stat* 2, 97-106.
- Becher, P.G., Bengtsson, M., Hansson, B.S., and Witzgall, P. (2010). Flying the fly: Long-range flight behavior of *Drosophila melanogaster* to attractive odors. *J Chem Ecol* 36, 599-607.
- Larsson, M.C., Domingos, A.I., Jones, W.D., Chiappe, M.E., Amrein, H., and Vosshall, L.B. (2004). Or83b encodes a broadly expressed odorant receptor essential for *Drosophila* olfaction. *Neuron* 43, 703-714.
- Maimon, G., Straw, A.D., and Dickinson, M.H. (2010). Active flight increases the gain of visual motion processing in *Drosophila*. *Nat Neurosci* 13, 393-399.
- Reiser, M.B., and Dickinson, M.H. (2008). A modular display system for insect behavioral neuroscience. *J Neurosci Methods* 167, 127-139.
- Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., *et al.* (2012). Fiji: An open-source platform for biological-image analysis. *Nat Methods* 9, 676-682.
- Tammero, L.F., Frye, M.A., and Dickinson, M.H. (2004). Spatial organization of visuomotor reflexes in *Drosophila*. *J Exp Biol* 207, 113-122.
- Zhu, J., Park, K.C., and Baker, T.C. (2003). Identification of odors from overripe mango that attract vinegar flies, *Drosophila melanogaster*. *J Chem Ecol* 29, 899-909.