SI Appendix: Full Description of Data Analysis Methods

Behavioral Data Analysis

Results from data analyses were expressed as mean ± standard error of the mean. Analyses of behavioral data were performed with GraphPad Prism (GraphPad Software, Inc., San Diego, California), and made use of two-way or one-way ANOVAs (with Bonferroni or Newman-Keuls post-hoc tests) and two-sample or one-sample t-tests. Sequential Bonferroni correction for multiple comparisons was performed with the Holm's method (1) whenever multiple independent t-tests were used in the same data set.

Two-Bottle Preference: All two-bottle preference tests were analyzed by calculating the preference ratios (P) for any particular stimulus as $P(stimulus1) = \frac{n(stimulus1)}{n(stimulus1) + n(stimulus2)}$, where n(stimulusX) denotes the total volume (rats) or the total number of licks (mice) for a particular stimulus X during a session. Average preference ratio across testing days was calculated for each animal (to account for side-bias). Note that in most cases, preference ratio was determined vs. water (i.e., stimulus 2 is water in most cases). Preference for each tastant was averaged across animals and expressed as mean ± SEM (standard error of the mean). Significance tests were based on one-sample t-tests against 0.5, which is the reference value meaning indifference with respect to stimulus 2. The concentration-preference functions for the ascending concentration series for nicotine and quinine in the rat were determined by fitting the means of each series to a sigmoidal dose-response function with

$P = \frac{0.5}{1+10^{n(b-x)}}$, where P is preference at a log[tastant]=x, b is the value of log[tastant] at half-maximal preference (i.e., logEC_{50}) and n is the slope factor. Given that the model was constructed with two-bottle preference values measured for two aversive tastants, it was constrained to a maximum value of 0.5 (maximal preference at concentrations below detection thresholds) and a minimum value of 0 (minimal preference at higher concentrations eliciting total rejection).

Rat Two-Alternative Choice Tests: Correct performance (i.e., proportion of correct responses) on all trials with a lever press was calculated across all tastants in a session and for each tastant across all concentrations.

Behavioral Analysis for the FR5 Schedule: For each stimulus in each test session, the amount of time elapsed between consecutive stimulus deliveries was measured. This span of time is referred to as the inter-tastant interval, or ITI. Average ITI's for each tastant in each animal (across all concentrations and testing days) were used to calculate the mean ± SEM tastant ITI across all animals. These were compared to establish relative preferences (more preferred tastants will elicit higher licking rates and thus shorter ITI's than less preferred ones with lower licking rates and long ITI's). Since this is an absolute measure averaged across a variable number of exposures to multiple concentrations of each tastant, the weighted average exposure concentration for each tastant ($[\overline{T}]$) was calculated such that $[\overline{T}] = \frac{\sum [C] * n_C}{\sum n_C}$, where $[C]$ is each concentration that was used and n_C is the number of exposures at each particular concentration.

Brief access tests: Lick ratios (LR) were defined as the amount of a particular stimulus consumed with respect to water, i.e., $LR(\text{stimulus}) = \frac{n(\text{stimulus})}{n(\text{water})}$, where n(X) denotes the total number of licks for a given stimulus X during a session. In animals tested with nicotine in the presence or absence of mecamylamine, lick ratios from single behavioral sessions were compared (mecamylamine session either the previous or the subsequent day to the nicotine only session – order counterbalanced across animals). In all other cases, lick ratios for each stimulus were averaged across days in each animal. Only data from sessions where animals gave 20 or more licks for water were analyzed. Thus, 3 WT mice (2 from the nicotine/quinine test; 1 from the nicotine/mecamylamine test) were excluded from analysis. Lick ratios for each tastant were averaged across animals and expressed as mean ± SEM. Average lick ratios for each experimental group were tested against 1.0, which is the reference value meaning indifference with respect to water.

Chorda Tympani Taste Nerve Recording Analysis

Data analysis pertaining to CT recordings was performed as described previously (2, 3). Phasic responses were obtained from the peak CT response while tonic responses were derived from the area under the curve for the last 30 seconds of the quasi-steady-state part of the response. In both cases responses were normalized to the tonic response to 300mM NH_4Cl. The responses thus quantified for each test stimulus in each animal were then averaged across 3 animals and expressed as mean ± SD (standard deviation). Analyses were performed using Sigmaplot (Systat Software Inc., San Jose, California) and with GraphPad Prism (GraphPad Software, Inc., San Diego, California), and made

use of two-way or one-way ANOVAs (with Bonferroni post-hoc tests) and two-sample t-tests.

Nicotine Dose-Response Curves and Effect of 0.3mm Mecamylamine: The data points for each nicotine-concentration series curve (with or without mecamylamine) and each animal group (rats, WT mice or KO mice) were fit to the Hill equation, i.e. $R = \frac{R_m c^n}{K^n + c^n}$, where *R* is the CT response at nicotine concentration *c*, R_m is the maximum response, *K* is the nicotine concentration that gives half-maximal response, and *n* is a number greater or equal to one. In each group, the asymptotic percent inhibition due to 0.3mM mecamylamine was calculated as: $100\frac{R_m(\text{control}) - R_m(\text{mecamylamine})}{R_m(\text{control})}$. The asymptotic percent inhibition in KO compared to WT was calculated as: $100\frac{R_m(\text{WT}) - R_m(KO)}{R_m(WT)}$.

Mecamylamine Dose-Inhibition Curves: The effects of mecamylamine on the responses to 10mM nicotine were plotted on a log scale of the molar concentration of mecamylamine. To avoid the minus infinity problem with the log of 0, the point corresponding to log[mecamylamine]=-5 is actually the response to 10mM nicotine alone (i.e, 0mM mecamylamine). A mecamylamine dose-inhibition curve was plotted according to $R = \frac{R_m}{1 + 10^{n(x-b)}} + R_a$, with R_m, b, R_a and n as the parameters to be fitted. R is the CT response at a log[mecamylamine]=x, R_m is the maximum response above the asymptotic response value (R_a) and b is the value of log[mecamylamine] for which

$R - R_a = 0.5R_m$ (i.e., $\log EC_{50}$). n is a fitted parameter associated to the interaction between mecamylamine and the nicotinic receptor.

Gustatory Cortex Neuronal Data Analysis

GC neuronal data analysis was conducted according to previously described methodology and will be described only briefly (4, 5). A total of 12 ensembles, each with 3 to 16 neurons, were analyzed separately. Given our previous findings (5), a 150ms window was taken from the fifth, reinforced lick and, for each neuron, spikes that fell within this window were binned in 15 ms increments. The spike counts corresponding to the third and seventh trials for each block of eight were dropped from the data set and the remaining data for each ensemble was analyzed with a Bayesian generalized linear model (GLM) (6). The reserved data was used to conduct the single trial predictions. In sessions where mecamylamine was added to stimuli, all trials with stimulus + mecamylamine were additionally dropped and used for predictions (see below).

Ensemble Modeling: Y_{ijkl} was defined as the number of spikes for neuron i, stimulus j, trial k, and time bin l while W_{ijk} is the number of spikes for neuron i, stimulus j, and trial k across the 150 ms window. Spikes were modeled as a Poisson distribution such that $Y_{ijkl} \sim Poi(\lambda_{ijkl})$ and $W_{ijk} \sim Poi(\lambda_{ijk})$ where λ_{ijkl} represents the Poisson distribution of the number of spikes in a 15 ms bin and $\lambda_{ijk} = \sum_l \lambda_{ijkl}$. Y_{ijkl}'s and W_{ijk}'s were assumed to be mutually independent. In this model $\ln(\lambda_{ijk}) = c_{ij} + d_{ijk}$, where $\exp(c_{ij})$ represents the mean firing rate of neuron i to stimulus j and d_{ijk} is an adjustment parameter across trials. For each neuron, the variability of the c_{ij}'s as j ranges over all stimuli captures how that neuron

changes its firing rate in response to different stimuli. Thus, λ_{ijk} represents the rate parameter of the model. Additionally, each spike in a 150ms window must fall into one of ten 15ms bins. The probability of falling in bin "l" is $a_{ijkl} = \lambda_{ijkl} / \lambda_{ijk}$. Thus, a_{ijkl} represents the temporal parameter of the model. Priors for c_{ij}, d_{ijk} and a_{ijkl} were set according to previous specifications (4). The terms c_{ij} and d_{ijk} were both censored for values below -5. A total of 40,000 iterations were conducted, and the first 1,000 samples were discarded, although the model always converged well before the first 1,000 iterations. Thirty-nine thousand iterations remained, and the data was then thinned in increments of 100 to make the analysis more computationally tractable. This resulted in a final set of 390 iterations from which to construct the posterior distributions for these terms.

Ensemble Data Prediction: When the model had run to completion, the distributions of the spike sums, $\hat{W}_{hij}$, was given by $\hat{W}_{hij} = Exp\left(\hat{c}_{hij} + \hat{d}_h\right)$, with h as the iteration number. The distributions of the ensemble firing patterns for the different stimuli, $\hat{Y}$, across all trials is then $\left(\hat{Y}_{hij1},...,\hat{Y}_{hij10}\right) = \hat{W}_{hij} * \left(\hat{a}_{hij1},...,\hat{a}_{hij10}\right)$. $\hat{a}$ reflects how the spikes are distributed in the bins and not the actual spike counts in each bin per se (the values of $\hat{a}$ sum to one across the time bins). The withheld data for a given ensemble and for a given stimulus in each bin was denoted as X_{ijkl}. A combined model, where $\hat{Y}$ is a function of both $\hat{W}$, the rate parameter, and $\hat{a}$, the temporal parameter, was used to calculate the probability of observing X given that a single delivery of a particular stimulus S_j had occurred ($P(X_{ijkl}|S_j)$). The stimulus that is predicted to have occurred given the ensemble firing pattern on the particular trial is the one for which $P(X_{ijkl}|S_j)$ is the greatest. These calculations are then repeated for each delivery of X.

Stimulus Predictions Across all Ensembles: To quantify prediction accuracy for each stimulus across all ensembles, the total number of correct predictions was divided by the total number of observations for the reserved data for that particular stimulus across all 12 ensembles. In sessions where mecamylamine was added to stimuli, trials with stimulus + mecamylamine were not included for this general assessment of GC ensemble stimulus/tastant predictions. Except where noted, we restricted our analyses to nicotine and quinine. For ensembles in which only one concentration of nicotine and quinine were tested (n = 3), the odds of correctly identifying stimulus identity by chance were one in two, or 50%. For ensembles tested with two concentrations of each tastant (n = 9), the odds of correctly identifying stimuli by chance were one in four, or 25%. Thus a weighted average of the chance level for each stimulus, given the number of stimuli tested and respective number of observations in each ensemble, was calculated such that $\overline{Ch}(S) = \frac{\sum_N Ch_N * t(S)_N}{\sum_N t(S)_N}$, with $\overline{Ch}$ representing the weighted chance level, S being the stimulus, Ch the chance level in ensemble N and t(S) the number of observed trials for the stimulus in ensemble N. The chance cutoff level is analogous to an α=0.05 (6, 7). Hence, if the ensemble prediction level is greater than the weighted chance level, then the ensembles collectively can discriminate that particular stimulus. Because these models are Bayesian, and each possesses a large number of parameters, it is standard practice to compare such models simply in terms of the percentage of data that each can predict (6, 7).

Tastant Predictions Across all Ensembles: The percentage of times that a particular tastant type (i.e. nicotine or quinine) was correctly identified was also determined. In this case, regardless of concentration, the total number of correct tastant classifications was divided by the total number of observations of that tastant across all concentrations. Since there were only two categories, the chance level was 50%. The tastant prediction rates were compared to the chance levels to determine whether tastant type could correctly classified by each model.

Predictions by Individual GC Ensembles: For each stimulus in each ensemble, the number of correct predictions was divided by the total number of observations for that particular stimulus to obtain the percent correct predictions. An ensemble was considered to correctly predict a particular stimulus whenever the percent correct predictions were above chance level for that ensemble (25 or 50%, see above). The proportion of correctly predicted stimuli was used as a measure of ensemble efficacy. To verify the importance of ensemble size in stimulus discrimination, the existence of a linear correlation between ensemble efficacy and ensemble size (number of neurons) was calculated. Since, under these conditions, the proportion of accurately predicted stimuli is on a discrete numerical scale (i.e, can only assume five values - 0, 0.25, 0.5, 0.75 or 1) the non-parametric Spearman's rho correlation coefficient was calculated.

Effects of mecamylamine on GC ensemble tastant prediction: In mecamylamine experiments, all trials for each stimulus to which mecamylamine had been added were dropped from the data set and used for additional analyses described here. Single trial

tastant identity predictions (nicotine vs. quinine – chance level of 50%) were conducted separately for the reserved data collected prior to and after mecamylamine addition. A comparison between the two allowed for the verification of the effects of peripheral nAChR blockade on ensemble prediction of tastants. To control for any unspecific effects of time or repeated exposure to nicotine, sessions with a similar number of total tastant blocks were chosen (n=4) and reanalyzed analogously to the mecamylamine sessions. Thus, GLMs were constructed excluding nicotine and water trials from the second half of the sessions and tastant identity predictions (nicotine vs. quinine) for nicotine trials were performed separately for the first half and second half of each session. Finally, tastant identity predictions for nicotine were repeated for the second half of both the mecamylamine and control sessions but a third classification possibility (water) was included. This last step was used to ascertain if the addition of mecamylamine introduced a bias in the classification of incorrectly predicted nicotine trials, different from that existing in control sessions. To that effect, the percentage of all incorrectly predicted nicotine trials classified as quinine was calculated. This number was compared with the chance level for a quinine error. Since some sessions were conducted with only one concentration of quinine (chance of quinine error of 50%) and others with two concentrations of quinine (chance of quinine error of 66.6%), the weighted chance level for quinine error was calculated for both the mecamylamine and control sessions, as described for the calculation of weighted chance levels for stimulus predictions.

References

1. Holm S (1979) A simple sequential rejective multiple test procedure. *Scand J Statistics* 6:65-70.
2. Lyall V*, et al.* (2005) Ethanol modulates the VR-1 variant amiloride-insensitive salt taste receptor. II. Effect on chorda tympani salt responses. *J Gen Physiol* 125(6):587-600.
3. Lyall V*, et al.* (2007) Effect of nicotine on chorda tympani responses to salty and sour stimuli. *J Neurophysiol* 98(3):1662-1674.
4. Stapleton JR, Lavine ML, Nicolelis MA, & Simon SA (2007) Ensembles of gustatory cortical neurons anticipate and discriminate between tastants in a single lick. *Frontiers in Neuroscience* 1(1):161-174.
5. Stapleton JR, Lavine ML, Wolpert RL, Nicolelis MA, & Simon SA (2006) Rapid taste responses in the gustatory cortex during licking. *J Neurosci* 26(15):4126-4138.
6. Dobson AJ (2002) *An introduction to generalized linear models* (Chapman and Hall/CRC, Boca Raton) second Ed.
7. Casella G & Berger RI (1987) Reconciling bayesian and frequentist evidence in the one-sided testing problem. *Journal of the American Statistical Association* 82(397):106-111.