Cortical recruitment determines learning dynamics and strategy

Sebastian Ceballo[1,3], Jacques Bourg[1,3], Alexandre Kempf[1,3], Zuzanna Piwkowska[1,4], Aurélie Daret[1], Thomas Deneux[1], Simon Rumpel[2] and Brice Bathellier[1,5]

[1] Unité de Neuroscience, Information et Complexité (UNIC), FRE 3693, Université Paris-Saclay and Centre National de la Recherche Scientifique, Gif-sur-Yvette, F-91198, France

[2] Institute of Physiology, Focus Program Translational Neuroscience, University Medical Center, Johannes Gutenberg University, Mainz, Germany.

[3] These authors contributed equally

[4] Present address: Institut Pasteur, Dynamic Neuronal Imaging Unit, Paris, France.

[5] Corresponding author: brice.bathellier@unic.cnrs-gif.fr, phone: +33(0)169823408

Abstract. Salience is a broad and widely used concept in neuroscience whose neuronal correlates, however, remain elusive. In behavioral conditioning, salience is used to explain various effects, such as stimulus overshadowing, and refers to how fast and strongly a stimulus is associated to a conditioned event. Here, we show that sounds of diverse quality, but equally intense and perceptually detectable, recruit surprisingly different levels of population activity in mouse auditory cortex. When using these sounds as cues in a Go/NoGo discrimination task, the degree of cortical recruitment matches the salience parameter of a reinforcement learning model used to analyze learning speed. Moreover, we confirm a generic prediction of the model by training mice to discriminate light-sculpted optogenetic activity patterns in auditory cortex verifying that cortical recruitment causally determines association or overshadowing of the stimulus components. This demonstrates that cortical recruitment captures major aspects of stimulus salience during reinforcement learning.

Main text

Introduction

Sensory stimuli can vary substantially in their efficacy to serve as a conditioned stimulus during behavioral conditioning. In the context of classical conditioning, a well-known example is the so called “overshadowing” effect: When animals are trained to associate two simultaneously presented stimuli to a specific unconditioned stimulus (e.g. foot-shock), it is often observed that, after training, the animal is conditioned more strongly to one stimulus than to the other (Kamin, 1965, 1969) indicating that one representation is favored over the other. In their seminal theoretical work originally developed for classical conditioning, but later extended to operand conditioning (e.g. in reinforcement learning), Rescorla and Wagner (Rescorla and Wagner, 1972) introduced the notion of salience to explain the overshadowing phenomenon. In their model, salience is a parameter affecting the speed at which a representation is associated to an action. Thus, when behavior reaches plateau performance and learning stops, the more salient of the two stimulus representations has built a stronger connection to the conditioned action, leading to overshadowing. While this theory captures a number of phenomena and provides the groundwork for important frameworks such as reinforcement learning (Dayan and Balleine, 2002; Sutton and Barto, 1998), the neural underpinnings of the salience parameter remain elusive.

Salience in this context is usually seen as the global amount of neural activity representing the stimulus, like in “pop out” models of attentional salience (Desimone and Duncan, 1995; Itti and Koch, 2001; Treue, 2003). This intuitively follows from the idea that if more spikes are involved in representing a stimulus, they can produce more synaptic weight changes, as expected from the firing rate sensitivity of typical learning rules (Bienenstock et al., 1982; Dayan and Abbott, 2001; Deneve et al., 2017; Fremaux and Gerstner, 2015; Martin et al., 2000), and thus modulate more rapidly the relevant connections. However widespread, this

idea lacks direct causal experimental verification in a learning task. Moreover, other theories propose that salience could also be encoded in other parameters such as neuronal synchrony levels (Fries, 2015; Shamma et al., 2011; VanRullen, 2003; Womelsdorf and Fries, 2007), which could influence learning via the temporal properties of biological learning rules (Buschman et al., 2012; Cassenaer and Laurent, 2012; Gutig and Sompolinsky, 2006; Kopell et al., 2014; Lee et al., 2009). Thus, the neuronal correlate of stimulus salience is a key question with broad implications for learning theories.

Salience is often associated to the physical magnitude of a stimulus, and thereby, salience coding can be confounded with intensity coding. However, the concept of salience can even be applied to stimuli of different quality (e.g. a sound *vs* a flash, a pure tone *vs* a white noise). Here, we identified sounds of equal physical intensity that recruit different global levels of auditory cortex activity in mice to investigate the neuronal correlate of stimulus salience independently of magnitude. Using a auditory discrimination task design that allows comparing the salience of two sounds independent of their perceptual similarity, we show that cortical recruitment impacts on the learning dynamics (Bathellier et al., 2013; Gallistel et al., 2004) in a manner similar to the salience parameter of a reinforcement learning model. Further, training mice to discriminate optogenetically triggered patterns that recruit different levels of cortical activity, we experimentally demonstrate that cortical recruitment determines which part of a compound stimulus drives a learned association while “overshadowing” other parts of the stimulus. This validates a generic prediction of the reinforcement learning model and causally establish the role of cortical recruitment as a neuronal parameter capturing stimulus salience.

Results

Sounds with identical physical levels and perceptual detectability elicit strikingly different levels of cortical recruitment.

It is often assumed, as a first approximation, that i) the physical intensity of a stimulus determines ii) the amount of recruited neuronal activity in the brain in as much as iii) its perceptual detectability, which is often used to describe the salience of a particular stimulus (Schnupp et al., 2011). We thus wondered how tightly these three properties are related for sounds, in light of the nonlinearities reported for the encoding of sound features (Chambers et al., 2014; Deneux et al., 2016). We chose three short complex sounds of 70ms duration containing a large range of frequencies and temporal modulations (**Figure 1A**) but normalized so that their mean sound pressure levels (standard deviation of the waveform) are equal. We first tested whether they are equally detectable by mice in a behavioral task. We initially trained mice to lick on a water spout after presentation of each of the three sounds (all mice experienced the three sounds in the same task) in order to obtain a reward. We then measured behavioral response probability to decreasing intensity levels of the sounds. We observed that, for all sounds, response probability steadily decreased down to chance level as measured in the absence of sound (**Figure 1B**). Yet, no significant difference in response probability curves was observed across the three sounds (Friedman test, $p = 0.43$, $n=6$), indicating that, in this context, detectability was mainly related to sound intensity and not quality.

We next wondered whether recruitment of neural activity in the auditory cortex in response to these three sounds was also independent of sound quality and essentially determined by the sound level. To do so, we used two-photon calcium imaging in awake, passively listening mice, a technique that offers access to large samples of neuronal activity. We imaged 6 mice which chronically expressed the GCAMP6s calcium indicator in auditory cortex after

injections with rAAV1 (**Figure 1C**). Recordings were followed by an automated image registration and segmentation algorithm (**Figure 1D**) (Roland et al., 2017) that allowed the isolation of 15,511 neurons across 27 imaging sites. The field-of-views were either 0.5×0.5 or 1×1 mm (**Figures 1C, D**), allowing for a rapid tiling of the full extent of primary and secondary auditory cortex. Cortical depths were randomly chosen ranging between 100 and 300 µm corresponding to layer 2/3. The mouse auditory cortex (primary + secondary) contains approximately 200,000 neurons in one hemisphere (Herculano-Houzel et al., 2013) and thus about 50,000 neurons in layer 2/3, so we expect our sample of ~15,000 neurons to be dense and representative for supragranular auditory cortex. Comparing the amplitude of the mean deconvolved calcium signals recorded across the entire duration of the response, we observed that, at 73 dB intensity, sound A elicited at least two-fold less cortical activity than sounds B and C (0.05 against 0.10 and 0.12 % $\Delta F/F.s^{-1}$, **Figure 1E**) and was consistently observed across mice (**Figure S1**). Furthermore, Sound A triggered a significant response in only about 35% of all neurons, while sounds B and C significantly excited 41% and 45% of all neurons (significance assessed with the Wilcoxon signed rank test, $p < 0.05$, **Figure 1E**). Furthermore, this difference in population responsiveness was consistent with previous, independent measurements performed under anesthesia (**Figure S1**) (Bathellier et al., 2012). In summary, our recordings show that despite similar levels of behavioral detectability and intensity level, sounds of different quality can recruit significantly different amounts of activity in the auditory cortex. It is noteworthy that all three sounds elicited distinct response patterns as evaluated by correlation-based population similarity measures. Thus sound identity could be decoded with high accuracy, on a single-trial basis, using linear classifiers (**Figure 1F**). This shows that sound discriminability was not affected by cortical recruitment.

Another discrepancy between cortical recruitment and the physical intensity of a stimulus can be observed using sounds with different temporal intensity profiles. Up-ramping sounds elicit

larger cortical responses in mice (Deneux et al., 2016) and other animals (Ghazanfar et al., 2002; Wang et al., 2014) than their time-symmetric down–ramps, despite the equality of their overall physical energies. This effect correlates with asymmetries in subjectively perceived loudness in humans (Neuhoff, 1998; Susini et al., 2007). We confirmed this result (**Figure 1G**) for 2s white noise sounds ramping between 60 and 85dB (mean deconvolved ΔF/F measured from 0 to 2.5 s after sound onset, up: 0.166 ± 0.58 ‰ vs down: 0.157 ± 0.57 ‰, Wilcoxon rank-signed test, p =0.034), with a clear effect even at the onset despite the lower start intensity level in up-ramps (mean between 0 and 0.5s for up: 0.80 ± 0.08 ‰ vs down: 0.62 ± 0.11 ‰, Wilcoxon rank-signed test $p = 3.75\times10^{-4}$, n = 29 sites, in 12 mice).

In summary, although intensity is obviously crucial for detectability, when sounds are played above the detection threshold, the specific amount of recruited cortical activity strongly depends on factors other than intensity and can vary across different sounds. Intrigued by this observation, we wondered whether cortical recruitment could be related to stimulus salience in a learning task.

Cortical recruitment controls specific aspects of learning speed.

To address this question, we aimed at behaviorally comparing the salience of pairs of sounds in a learning task. Classically, this is done using an overshadowing paradigm in which two stimuli (e.g. a flash and a sound) are conditioned together, as a compound stimulus, to an unconditioned stimulus (e.g. foot-shock). Then, salience is derived from the level of the conditional response elicited by each stimulus component individually. While this approach is valid when the compound is made of stimuli from two different sensory modalities, two simultaneous sounds are likely to fuse perceptually, preventing to address their individual salience with the classical overshadowing design (Brandon et al., 2000). Thus, we measured stimulus salience through learning speed, using an auditory-cued Go/No-Go task. To do so,

water-deprived mice were first trained to visit a lick-port and to receive a water reward if they licked at a water spout after being presented with a S+ sound. This pre-training phase mainly aimed at raising motivation in all mice and was not used to measure learning speed. When mice collected rewards in at least 80% of their port visits, the Go/NoGo task was started by introducing a non-rewarded S- sound in half of the trials (**Figure 2A**). After a large number of trials, mice succeeded to both sustain licking to the S+ and withdraw from licking for the S- (**Figure 2B,C**), thereby demonstrating their ability to discriminate the two sounds. Importantly, as typically observed in such tasks {Bathellier, 2013 #946}, the S+ sound was rapidly associated to the lick response and the rate limiting factor in the acquisition of the task was to associate the S- sound with correct rejections (**Figure 2C**). Hence, learning speed depends more on the salience of the S- than of the S+ sound in this task. Salience of two stimuli X and Y can thus be compared irrespective of their discriminability by comparing the learning speed of the X versus Y Go/NoGo discrimination when X is S- against the speed observed when Y is S-. We therefore trained six cohorts of mice to compare the salience of sounds pairs A-B, A-C and B-C. Plotting the population learning curves for each cohort and each sound pair, we strikingly observed that the average learning speed was systematically slower when cortical recruitment for the S- sound was smaller than for the S+ sound (**Figure 2B**). This suggested a relation between cortical recruitment and sound salience, and corroborated an observation made in an earlier study (Deneux et al., 2016) for the discrimination of the up- and down-ramping white noise ramps shown in **Figure 1F**.

However, looking at learning curves from individual mice, we noticed that the qualitative difference observed at the group level hides a more complex effect. As often observed in animal training (Gallistel et al., 2004) and as we previously reported for the particular task used in this study (Bathellier et al., 2013), most individual learning curves had a sigmoidal rather than exponential time course. Specifically, the curves displayed a delay phase with no

increase in performance followed by a learning phase with an often very steep performance increase. Also, the duration of each phase was highly variable across animals as exemplified in **Figure 2C**. We wondered whether cortical recruitment was affecting one particular phase or both. We performed a sigmoidal fit on each individual curve from which the delay phase duration was computed as the number of trials necessary to reach 20% of maximal performance level, and learning phase duration as the number of trials necessary to go from 20% to 80% maximal performance. These measurements showed for each sound pair tested that the learning phase was significantly longer when the S- sound recruited less cortical activity than the S+ sound, (**Figure 2D**, Friedman test, $p = 0.016$). On the contrary, we could not detect any correlation between the level of cortical recruitment and the delay phase (**Figure 2D**, Friedman test, $p = 0.45$). In addition, we noticed that cortical recruitment had an effect not only on the mean duration of the learning phase at the group level but also on the inter-individual variability. Cohorts having an S- sound recruiting more activity than the S+ sound had a relatively homogenous learning phase duration, especially for the two sound pairs with a large difference of cortical recruitment, while the cohort with the opposite sound assignment had some animals with a fast and some with a much slower learning phase (**Figure 2E**). Finally, re-analyzing in detail the learning curves obtained in our previous study for the discrimination of up- and down-ramping sounds (Deneux et al., 2016), we observed the same type of increase in mean learning rate duration and variability when the S- sound was the down-ramp (**Figure 2F**).

Thus altogether, in all four Go/NoGo discrimination tasks analyzed, we observed that mice display a much longer learning phase when the S- sound recruits less cortical activity than the S+ sound as compared to animals with the opposite sound assignment. This indicates that cortical recruitment controls specific parameters of learning dynamics.

A reinforcement learning model captures the effects of cortical recruitment on learning dynamics

To analyze in detail the effect of cortical recruitment on learning, we employed a recently developed model of the discrimination task based on the reinforcement learning framework extending the Rescorla-Wagner model to a simple but more biologically interpretable neural model (**Figure 3A**) (Bathellier et al., 2013). In short, the model postulates that the associative learning occurs by adjusting the synaptic weights between "sensory" and "decision" neural populations described by population firing rate variables. At the input, two populations are specific for the S+ and S- sounds respectively, which we denote as $\hat{S}$+ and $\hat{S}$-, and one population, $\hat{C}$, which represents information common to S+ and S- trials (e.g. overlap between the S+ and S- representations or activity independent of sound, for example, related to visiting the lick port). Population $\hat{C}$ is an essential element of the model to reproduce high initial 'hit'-rates combined with delayed discrimination learning (Bathellier et al., 2013). The "decision" population has two ensembles: one promoting and one inhibiting licking (**Figure 3A**). Adjustment of synaptic weights happens through a Hebbian learning rule modulated by Rescorla and Wagner's (1972) δ-rule which conditions weight updates to a non-zero reward expectation error. However, the δ-rule is asymmetric, meaning that the learning rate is larger by a factor v when an unexpected reward occurs, as compared to when an expected reward does not occur. This asymmetry is crucial for capturing the asymmetry of the learning process (i.e. the fast adjustment of the 'hit'-rate, and slower adjustment of the 'correct rejection'-rate; **Figures 2C and 3D**). In addition, synaptic updates are multiplicative meaning that weight updates are proportional to the current weight value (Gutig et al., 2003; Koulakov et al., 2009; Loewenstein et al., 2011; Morrison et al., 2007). The multiplicative learning rule is crucial to

capture delay and leaning phases observed in behavior (**Figures 2C and 3D**). Furthermore, this feature makes learning speed sensitive to the initial synaptic weights, which become important parameters that can even account for most of the interindividual variability (Bathellier et al., 2013). Most importantly here, the model has an explicit salience parameter, that corresponds to the activity of the $\hat{S}$+ and $\hat{S}$- populations, which makes it well suited to interpret our experiments.

We therefore wondered whether this model reproduces the observed effects of cortical recruitment on mean learning speed and its variability. To do so, we doubled the firing rate of either the $\hat{S}$+ or $\hat{S}$- sub-populations to model differences in cortical recruitment. In our model, it is important to run the simulations within a realistic range of initial synaptic weight parameters, since they are a major determinant of the dynamics of the learning process. Thus, we ran simulations with 15 different sets of initial synaptic weight parameters obtained in a previous set of experiments (Bathellier et al., 2013) by fitting the individual learning curves from 15 mice trained to discriminate sounds A and B in a task identical to the one used in this study. We observed that when the S- recruits less activity than the S+, both the mean learning phase duration and its variability increased compared to when the S- recruits more activity than the S+, which strikingly resembled our experimental observations (n = 15 parameter sets, Kolmogorov-Smirnov test comparing the two distributions, $p = 8\times10^{-7}$, **Figure 3B**). Again, no effect was observed for the delay phase (n = 15 parameter sets, Kolmogorov-Smirnov test, p = 0.13, **Figure 3B**).

To understand the origin of this phenomenon in our simulations, we plotted learning phase duration against the values of the three initial connection weight parameters (**Figure 3C**, w_{E_com} and w_{I_com} for the $\hat{C}$ population, and a single initial value $w_{I|E_sounds}$ for the four weights of the $\hat{S}$+ and $\hat{S}$- population). When S- recruited more activity than S+, initial weights had no

systematic influence on learning phase duration (**Figure 3C**). When S- recruited more activity than S+, there was also no influence of the initial synaptic weights from the S+ and S- sub-populations (Pearson correlation coefficient ρ = -0.09, p = 0.74, n=15) but learning phase duration was negatively correlated to the initial weights from the "common" input population to both the excitatory (ρ = -0.60, p = 0.016, n=15) and the inhibitory (ρ = -0.68, p = 0.005, n=15) decision populations. This negative correlation was due to a specific increase in learning phase duration when S- recruits less activity than S+ and when the "common" population $\hat{C}$ is initially weakly connected to decision (**Figure 3C, D and 4A**). This specific increase could be observed only when the learning rule was asymmetric (i.e. $v > 1$) and the effect was stronger for large asymmetry values (**Figure 4A**). The reason can be understood intuitively. When the learning rule is symmetric ($v = 1$), learning dynamics for S+ and S- stimuli are symmetric and therefore the overall dynamics is expected to remain the same after swapping the two stimuli.

Further simulation involving a systematic analysis of initial synaptic weight contributions also showed that low initial weights to the inhibitory rather than to the excitatory decision population, is the determining factor for the observed increase in learning phase duration as compared to other initial weights (**Figure 4B-D**). Altogether, our simulations show that differences in neuronal recruitment between two sensory stimuli can specifically impact the learning dynamics of their discrimination, however this effect depends on the initial states of some of the synapses important in driving behavior. We therefore used the model to search for additional effects of the respective neuronal recruitment induced by the S+ and S- stimuli, that are independent of initial synaptic states and that could be tested in a well-controlled experimental setting.

Cortical recruitment determines learning strategy in the model through a mechanism related to overshadowing.

One important feature of our model is that there is redundant information representing the two discriminated stimuli (e.g. S+ is represented by populations $\hat{S}$+ and $\hat{C}$). Thus different final configurations of connections can solve the discrimination task with equal efficiency (see also Supplementary Note). To better understand which part of the sensory input representations actually drives learning and behavior in the model, we plotted the time-course of both excitatory and inhibitory connection weights (**Figure 5**). The plots revealed that the cues selected for driving behavior depend on the relative neuronal recruitment for the discriminated stimuli. When S- recruits more activity than S+, the cue for withdrawing from licking is unsurprisingly the $\hat{S}$- population, however, the lick response is not driven by the $\hat{S}$+ population but the $\hat{C}$ population, which represents all cues common to S+ and S- (**Figure 5A, B**). In contrast, when S- recruits less activity than S+, the cue selected for the lick response is the $\hat{S}$+ population, but the dominating cue for refraining from licking is now the $\hat{C}$ population (**Figure 5C, D**). These effects are closely related to overshadowing, as they concretely stem from the overshadowing of the "common" population by the most "salient" of the $\hat{S}$+ or $\hat{S}$- populations. Because of this overshadowing, the common $\hat{C}$ population drives the action opposite to the stimulus recruiting more activity. Interestingly, this effect, in contrast to the learning speed effects, appeared to be largely independent of initial synaptic weights (**Figure 5**) and thus represent a robust manifestation of the role of neuronal recruitment in discrimination learning. The reason of this effect is that the learning process compensates for differential neuronal recruitment by building stronger connections for the stimulus recruiting less activity, in order to get equal performance for both stimuli (see Supplemental Experimental Procedures for more details). Although $\hat{C}$ belongs to both stimuli, this

compensation process mechanically boosts its connections to the outcome corresponding to the weaker stimulus.

Most importantly, this provides a clear prediction that can be used to experimentally demonstrate that cortical recruitment corresponds to stimulus salience in our learning task: the behavioral response of the animal triggered by the activation of the common $\hat{C}$ population alone will reveal the selected learning strategy and will depend on the saliency of $\hat{S}$+ and $\hat{S}$- populations. More precisely, activation of the $\hat{C}$ population alone should drive licking when S- recruits more activity, and should not drive licking when S- recruits less activity (**Figure 5)**. As we could show analytically (see Supplemental Experimental Procedures), this prediction is a very general property of the reinforcement learning model independent of specific parameters. Hence, it accurately reflects the basic principles by which stimulus salience, as described in the model, impacts learning. We therefore looked for an opportunity to test this prediction experimentally.

Optogenetic demonstration of the role of cortical recruitment in selecting learning strategy.

Testing this prediction is, however, not straightforward in the context of the specific behavioral task design we had chosen for the initial series of experiments. In freely moving mice there are two possible confounding cues that are common to the S+ and S- sounds and that could drive the common ($\hat{C}$) population of neurons: (i) the overlap of their representations and (ii) all cues related to the decision to visit the lick port. Thus, it is difficult to identify which are exactly the common cues driving behavior, making it essentially impossible to activate them in isolation. Furthermore, given the complexity and diversity of neuronal tuning properties in the auditory cortex, it is difficult to design sound stimuli that either have no overlap in their representations or that would allow independent manipulation of the neuronal

populations forming the independent or common components of their neuronal representations.

In the light of these confounding factors, we decided to test the model's prediction in a partially artificial but better controlled experiment in which head-fixed mice had to discriminate optogenetically driven cortical ensembles. To do so we used a custom-made setup, based on a video projector (Dhawale et al., 2010), to project precise 2D light patterns through a cortical cranial window placed above the auditory cortex in Emx1-Cre x Ai27 mice expressing channelrhodopsin in excitatory cortical neurons (**Figure 6A and B**). Using intrinsic imaging, we identified the main tonotopic fields of mouse auditory cortex (**Figure 6C**) (Bathellier et al., 2012; Moczulska et al., 2013) and thereby reliably positioned optogenetic stimulation spots in homologous regions across mice (**Figure 6C**). We defined three circular optogenetic stimulation spots out of which we constructed two stimuli. One of the three spots, the $\hat{C}$ spot, was common to the two S+ and S- stimuli and the two other spots corresponded to the stimulus-specific neuronal populations $\hat{S}$+ and $\hat{S}$- (e.g. Figure 6B). Thus, the cue S+ used for Go-trials consisted of simultaneous activation of $\hat{S}$+ and $\hat{C}$ and the cue S- consisted of simultaneous activation of $\hat{S}$- and $\hat{C}$. Furthermore, we were able to exclude cues common to S+ and S-, related to the visit of the port because, in the head-fixed task design, mice did not initiate the trials, which occurred at random time interval. We thereby ensured that the only cue common to the S+ and S- stimuli was the $\hat{C}$ spot. We doubled the diameter of either the $\hat{S}$+ or the $\hat{S}$- spot to create a difference in cortical recruitment between the two input representations as in the model. Mice were then initially trained to obtain a water reward by licking after the coincident activation of the $\hat{S}$+ and $\hat{C}$ spot. When 80% performance was reached in this detection task, the discrimination training started. During this stage, activation of the $\hat{C}$ population continued to occur in all trials: in half of the trials together with $\hat{S}$+ and in the other half together with $\hat{S}$-. Mice kept licking in the presence of the S+ cue spot and

learned, within few a hundreds of trials to avoid licking in the presence of the S- cue (**Figure 6D**), reaching a mean steady state performance of 94.7% ± 4.5% correct trials (hit rate 93.9% ± 3.3%, false alarm rate 4.5% ± 1.3%, n=8 mice, see **Figure 6E&F**). Importantly, in this task setting, the stringent definition of the common Ĉ population, activated during the initial motivation training, likely resulted in the systematic establishment of strong initial connections for this population at the beginning of the discrimination training, leading to homogenous durations of the learning phase (see **Figure 6D**) independent of salience, as detailed in the simulations shown in **Figures 3-5**. Yet, the more general prediction of the model regarding learning strategy could be verified.

Indeed, once the mice had learned the behavioral task, we measured their response to Ĉ activation alone in occasional catch trials that were not rewarded (catch trial probability = 0.1; 15 catch trials per mouse). In the group of mice which had a larger Ŝ- spot, we observed that activation of Ĉ elicited strong licking responses (83% ± 10% response probability, n = 4 mice, **Figure 6E**). In contrast, in the group of mice that had a larger Ŝ+ spot, activation of Ĉ elicited no licking response (0% ± 0% licks per trials, n = 4 mice, **Figure 6F**). These results confirm the predictions of the model, and demonstrate, in a causal manner, that cortical recruitment affects associative learning. In particular, it shows that the internal sensory representations available for association to a particular action are selected according to the amount of activity they correspond to. In classical models of associative learning (Rescorla and Wagner, 1972), stimulus salience is the parameter that determines which one of two stimuli having the same meaning is chosen for association with a particular behavior (overshadowing effect). We here show that neuronal recruitment in the mouse auditory cortex plays the role of the salience parameter in the discrimination learning task.

Discussion.

Combining, behavioral measurements, large scale two-photon imaging, optogenetics and theoretical modeling we have shown that sounds of different quality can elicit highly variable levels of neuronal recruitment in auditory cortex, measured as the mean amount of activity in a representative subsample of neurons. We also showed that cortical recruitment levels are tightly correlated with learning speed effects, using a Go/NoGo behavioral task. These effects can be precisely reproduced by a reinforcement learning model of the task when cortical recruitment, as measured experimentally, corresponds to the stimulus salience parameter of the model. Finally, training mice to discriminate optogenetically engineered cortical patterns, and manipulating these patterns, we showed that neuronal recruitment determines which elements of the cortical representation are selected to be associated to each conditioned action. This corroborates in a causal manner the idea that cortical recruitment is a neuronal correlate of stimulus salience.

What is the origin of the differences in cortical recruitment between sounds that have equal intensity levels? In a recent study, we have suggested that nonlinearities of the auditory system select or amplify particular acoustic features over others such as the on- and offset features of up- and down-intensity ramps (**Figure 1G**) (Deneux et al., 2016). This process starts as early as the cochlea which is more sensitive to some frequencies than others (Fay, 1988; Henry, 1979) and further nonlinearities along the multiple stages of the auditory system also contribute (Nelken, 2004). Thus, the relative level of cortical recruitment observed for different sounds represents interesting information about central processes for acoustic feature extraction and selection. This suggests that cortical recruitment should at least partially reflect neuronal recruitment in higher subcortical areas such as thalamus and therefore should be a relevant measure of salience independent of whether the discrimination tasks requires auditory cortex (Kato et al., 2015; Kuchibhotla et al., 2017; Znamenskiy and Zador, 2013) or do not require it (Gimenez et al., 2015; Pai et al., 2011; Rybalko et al., 2006).

The idea that the amount of neuronal activity recruited by a stimulus influences behavior has been proposed in different contexts. For example, several studies indicate that attention can boost the activity of the neurons representing behaviorally relevant stimuli (Desimone and Duncan, 1995; Fritz et al., 2007; Goard and Dan, 2009) and thereby make it more discriminable from other stimuli (Zhang et al., 2014). Earlier reports, using direct microstimulation of the cortex as a stimulus, showed that low levels of neuronal recruitment can impact detection probability (Houweling and Brecht, 2008; Huber et al., 2008). In contrast here, we show a different impact of cortical recruitment for levels of evoked activity clearly above detection threshold. Namely, we show that the intrinsic difference in neuronal recruitment across stimuli, once they reach intensity levels beyond detection threshold, impact the learning process by which the stimuli can be associated to particular responses. This is well captured by our model which provides an extension and a simple neuronal implementation of the idea proposed by Rescorla and Wagner (Rescorla and Wagner, 1972). In the Rescorla-Wagner model, the effective associative learning rate varies across stimuli via a salience parameter. In our model, the effective learning rate depends on the product of synaptic strength (multiplicative rule) and neuronal recruitment in the sensory representation (Hebbian rule). The level of cortical recruitment experimentally assessed for each stimulus turned out as an excellent estimate for the salience parameter in the model. However, a large cortical recruitment alone is not sufficient to obtain rapid associative learning of a stimulus as the connections from the stimulus representation to the decision neurons must also be strong enough. As a consequence, the effect of neuronal salience on initial learning speed can be occluded by large inter-individual differences in initial synaptic weights (Bathellier et al., 2013). This probably explains why no overt effect of neuronal recruitment is observed in the duration of the delay phase (**Figure 2D-F**) which is mainly determined by initial synaptic weights in this task (Bathellier et al., 2013). The same mechanism also explains why neuronal

recruitment does not always impact the learning phase duration. Indeed, even if learning speed is proportional to the product of neuronal recruitment and connectivity, learning tends to compensates weak recruitment by stronger connections (see Supplemental Experimental Procedures), a phenomenon that tends to stabilize learning speed, except for particular initial conditions (**Figures 3-5**).

In fact, Rescorla and Wagner suggested that the main impact of the salience parameters in learning is the determination of which stimulus within a group of stimuli will be strongly associated to the behavioral output while others may be overshadowed (Rescorla and Wagner, 1972). In principle, this rule not only applies to groups of physically distinct stimuli such as light and sound, but also to components of the same physical stimulus (e.g. different frequency bands in a multi-frequency sound) or, rather, to distinct subsets of the neuronal representation of a stimulus (e.g. neurons representing different sound frequencies). There have been no experimental attempts to test the latter hypothesis, as the simultaneous presentation of different external stimuli (e.g. two pure tone frequencies, or a sound and vibration) does not necessarily lead to a neuronal representation that corresponds to the linear addition of the representation of each component alone. This leads to complex effects which cannot be explained by the Rescorla-Wagner model alone and would require intricate modeling of perceptual interactions (Brandon et al., 2000). Here, using conditioning to compound stimuli composed of multiple optogenetically activated neuronal ensembles of different sizes in auditory cortex, we disentangled the perceptual interaction between components of natural stimuli. We directly showed that the components of a unimodal, artificial sensory representation are associated to conditioned actions according to the respective amount of neuronal activity that they generate (**Figure 6E&F**). This result demonstrates that the generic principles leading to overshadowing in classical conditioning also hold for subsets of a unimodal stimulus representation. Moreover, these optogenetic

experiments causally demonstrate that cortical recruitment plays the role of Rescorla and Wagner's salience parameter for choosing the stimulus representation that will become associated to the conditioned response (**Figure 6E&F**). This indicates that when conditioned to complex stimuli, the brain will choose the elements of their representations which recruit the largest amount of activity for association with a particular behavioral output.

In our study, we investigated the relationship of cortical recruitment and salience in the context of a reinforcement task in the Go/NoGo design. However, cortical recruitment could play a role also in different contexts. For example, frequent exposure to auditory stimuli leads to a strong adaptation of auditory responses (Kato et al., 2015). This decrease in cortical recruitment could be related to latent inhibition, a phenomenon describing that the effectiveness of a stimulus to serve as a conditional stimulus decreases dramatically when it has been presented to the subject without reinforcement before conditioning (Lubow and Weiner, 2010). We therefore speculate that cortical recruitment could also serve as a useful parameter explaining learning phenomena in other paradigms.

Acknowledgements

We thank K. Kuchibhotla, E. Harrel and M. Stüttgen for comments on the manuscript, P. Pindi, S. Sikirić and L. François for help with behavioral and imaging experiments. We thank the GENIE Project, Janelia Farm Research Campus, Howard Hughes Medical Institute, for GCAMP6s constructs. This work was supported by the Agence Nationale pour la Recherche (ANR "SENSEMAKER"), the Fyssen foundation, the DIM "Region Ile de France", the Marie Curie Program (CIG 334581), the Internastional Human Frontier Science Program Organization (CDA-0064-2015), by the Fondation pour l'Audition (Laboratory grant), the École Doctorale Frontières du Vivant (FdV) – Programme Bettencourt (support to AK), the

DIM Cerveau et Pensée and Ecole des Neurosciences de Paris Ile-de-France (ENP, support to SC) and the Deutsche Forschungsgemeinschaft (DFG CRC1080/2).

Author contributions

BB and SR designed the study. AK, SC, and BB performed and analyzed the imaging experiments. BB and JB performed the modeling. AD and BB performed and analyzed behavioral experiments. ZP designed the patterned optogenetic setup and SC performed the optogenetic experiments. TD designed software for data analysis and behavior. BB and SR wrote the manuscript with comments from all authors.

CRediT Taxonomy: Conceptualization, B.B. and S.R.; Methodology, Z.P., B.B.; Software, T.D.; Investigation, S.C., A.K., J.B., A.D., B.B.; Formal Analysis: J.B., B.B.; Writing – Original Draft, B.B. and S.R.; Funding Acquisition, B.B. and S.R.;

Declaration of Interests

The authors declare no competing interests.

References

Bathellier, B., Tee, S.P., Hrovat, C., and Rumpel, S. (2013). A multiplicative reinforcement learning model capturing learning dynamics and interindividual variability in mice. Proc Natl Acad Sci U S A *110*, 19950-19955.

Bathellier, B., Ushakova, L., and Rumpel, S. (2012). Discrete neocortical dynamics predict behavioral categorization of sounds. Neuron *76*, 435-449.

Bienenstock, E.L., Cooper, L.N., and Munro, P.W. (1982). Theory for the development of neuron selectivity: orientation specificity and binocular interaction in visual cortex. J Neurosci *2*, 32-48.

Brandon, S.E., Vogel, E.H., and Wagner, A.R. (2000). A componential view of configural cues in generalization and discrimination in Pavlovian conditioning. Behav Brain Res *110*, 67-72.

Buschman, T.J., Denovellis, E.L., Diogo, C., Bullock, D., and Miller, E.K. (2012). Synchronous oscillatory neural ensembles for rules in the prefrontal cortex. Neuron *76*, 838-846.

Cassenaer, S., and Laurent, G. (2012). Conditional modulation of spike-timing-dependent plasticity for olfactory learning. Nature *482*, 47-52.

Chambers, A.R., Hancock, K.E., Sen, K., and Polley, D.B. (2014). Online stimulus optimization rapidly reveals multidimensional selectivity in auditory cortical neurons. J Neurosci *34*, 8963-8975.

Cohen, J.Y., Haesler, S., Vong, L., Lowell, B.B., and Uchida, N. (2012). Neuron-type-specific signals for reward and punishment in the ventral tegmental area. Nature *482*, 85-88.

Dayan, P., and Abbott, L.F. (2001). Theoretical neuroscience : computational and mathematical modeling of neural systems (Cambridge, Mass.: Massachusetts Institute of Technology Press).

Dayan, P., and Balleine, B.W. (2002). Reward, motivation, and reinforcement learning. Neuron *36*, 285-298.

Deneux, T., Kempf, A., Daret, A., Ponsot, E., and Bathellier, B. (2016). Temporal asymmetries in auditory coding and perception reflect multi-layered nonlinearities. Nature Communications *7*, 12682.

Deneve, S., Alemi, A., and Bourdoukan, R. (2017). The Brain as an Efficient and Robust Adaptive Learner. Neuron *94*, 969-977.

Desimone, R., and Duncan, J. (1995). Neural mechanisms of selective visual attention. Annu Rev Neurosci *18*, 193-222.

Dhawale, A.K., Hagiwara, A., Bhalla, U.S., Murthy, V.N., and Albeanu, D.F. (2010). Non-redundant odor coding by sister mitral cells revealed by light addressable glomeruli in the mouse. Nat Neurosci *13*, 1404-1412.

Fay, R.R. (1988). Comparative psychoacoustics. Hear Res *34*, 295-305.

Fremaux, N., and Gerstner, W. (2015). Neuromodulated Spike-Timing-Dependent Plasticity, and Theory of Three-Factor Learning Rules. Front Neural Circuits *9*, 85.

Fries, P. (2015). Rhythms for Cognition: Communication through Coherence. Neuron *88*, 220-235.

Fritz, J.B., Elhilali, M., David, S.V., and Shamma, S.A. (2007). Auditory attention--focusing the searchlight on sound. Curr Opin Neurobiol *17*, 437-455.

Gallistel, C.R., Fairhurst, S., and Balsam, P. (2004). The learning curve: implications of a quantitative analysis. Proc Natl Acad Sci U S A *101*, 13124-13131.

Ghazanfar, A.A., Neuhoff, J.G., and Logothetis, N.K. (2002). Auditory looming perception in rhesus monkeys. Proc Natl Acad Sci U S A *99*, 15755-15757.

Gimenez, T.L., Lorenc, M., and Jaramillo, S. (2015). Adaptive categorization of sound frequency does not require the auditory cortex in rats. J Neurophysiol *114*, 1137-1145.

Goard, M., and Dan, Y. (2009). Basal forebrain activation enhances cortical coding of natural scenes. Nat Neurosci *12*, 1444-1449.

Gutig, R., Aharonov, R., Rotter, S., and Sompolinsky, H. (2003). Learning input correlations through nonlinear temporally asymmetric Hebbian plasticity. J Neurosci *23*, 3697-3714.

Gutig, R., and Sompolinsky, H. (2006). The tempotron: a neuron that learns spike timing-based decisions. Nat Neurosci *9*, 420-428.

Henry, K.R. (1979). Auditory nerve and brain stem volume-conducted potentials evoked by pure-tone pips in the CBA/J laboratory mouse. Audiology *18*, 93-108.

Herculano-Houzel, S., Watson, C., and Paxinos, G. (2013). Distribution of neurons in functional areas of the mouse cerebral cortex reveals quantitatively different cortical zones. Front Neuroanat *7*, 35.

Houweling, A.R., and Brecht, M. (2008). Behavioural report of single neuron stimulation in somatosensory cortex. Nature *451*, 65-68.

Huber, D., Petreanu, L., Ghitani, N., Ranade, S., Hromadka, T., Mainen, Z., and Svoboda, K. (2008). Sparse optical microstimulation in barrel cortex drives learned behaviour in freely moving mice. Nature *451*, 61-64.

Itti, L., and Koch, C. (2001). Computational modelling of visual attention. Nat Rev Neurosci *2*, 194-203.

Kamin, L.J. (1965). Temporal and intensity characteristics of the conditioned stimulus. In Classical conditioning, P. W.F., ed. (New York: Appleton-Century-Crofts), pp. 118-147.

Kamin, L.J. (1969). Predictability, surprise, attention, and conditioning. In Punishment and Aversive Behavior, B.A. Campbell, and R.M. Church, eds. (New York: Appleton-Century-Crofts), pp. 279-296.

Kato, H.K., Gillet, S.N., and Isaacson, J.S. (2015). Flexible Sensory Representations in Auditory Cortex Driven by Behavioral Relevance. Neuron *88*, 1027-1039.

Kopell, N.J., Gritton, H.J., Whittington, M.A., and Kramer, M.A. (2014). Beyond the connectome: the dynome. Neuron *83*, 1319-1328.

Koulakov, A.A., Hromadka, T., and Zador, A.M. (2009). Correlated connectivity and the distribution of firing rates in the neocortex. J Neurosci *29*, 3685-3694.

Kuchibhotla, K.V., Gill, J.V., Lindsay, G.W., Papadoyannis, E.S., Field, R.E., Sten, T.A., Miller, K.D., and Froemke, R.C. (2017). Parallel processing by cortical inhibition enables context-dependent behavior. Nat Neurosci *20*, 62-71.

Lee, S., Sen, K., and Kopell, N. (2009). Cortical gamma rhythms modulate NMDAR-mediated spike timing dependent plasticity in a biophysical model. PLoS Comput Biol *5*, e1000602.

Loewenstein, Y., Kuras, A., and Rumpel, S. (2011). Multiplicative dynamics underlie the emergence of the log-normal distribution of spine sizes in the neocortex in vivo. J Neurosci *31*, 9481-9488.

Lubow, R.E., and Weiner, I. (2010). Latent inhibition: Data, theories, and applications to schizophrenia. (New York: Cambridge University Press).

Martin, S.J., Grimwood, P.D., and Morris, R.G. (2000). Synaptic plasticity and memory: an evaluation of the hypothesis. Annu Rev Neurosci *23*, 649-711.

Moczulska, K.E., Tinter-Thiede, J., Peter, M., Ushakova, L., Wernle, T., Bathellier, B., and Rumpel, S. (2013). Dynamics of dendritic spines in the mouse auditory cortex during memory formation and memory recall. Proc Natl Acad Sci U S A *110*, 18315-18320.

Morrison, A., Aertsen, A., and Diesmann, M. (2007). Spike-timing-dependent plasticity in balanced random networks. Neural Comput *19*, 1437-1467.

Nelken, I. (2004). Processing of complex stimuli and natural scenes in the auditory cortex. Curr Opin Neurobiol *14*, 474-480.

Neuhoff, J.G. (1998). Perceptual bias for rising tones. Nature *395*, 123-124.

Pai, S., Erlich, J.C., Kopec, C., and Brody, C.D. (2011). Minimal impairment in a rat model of duration discrimination following excitotoxic lesions of primary auditory and prefrontal cortices. Front Syst Neurosci *5*, 74.

Rescorla, R.A., and Wagner, A.R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In Classical Conditioning II: Current Research and Theory B. A.H., and P. W.F., eds. (New York: Appleton Century Crofts), pp. 64-99.

Roland, B., Deneux, T., Franks, K.M., Bathellier, B., and Fleischmann, A. (2017). Odor identity coding by distributed ensembles of neurons in the mouse olfactory cortex. Elife *6*.

Rybalko, N., Suta, D., Nwabueze-Ogbo, F., and Syka, J. (2006). Effect of auditory cortex lesions on the discrimination of frequency-modulated tones in rats. Eur J Neurosci *23*, 1614-1622.

Schnupp, J., Nelken, I., and King, A. (2011). Auditory neuroscience : making sense of sound (Cambridge, Mass.: MIT Press).

Schultz, W., Dayan, P., and Montague, P.R. (1997). A neural substrate of prediction and reward. Science *275*, 1593-1599.

Shamma, S.A., Elhilali, M., and Micheyl, C. (2011). Temporal coherence and attention in auditory scene analysis. Trends Neurosci *34*, 114-123.

Susini, P., McAdams, S., and Smith, B. (2007). Loudness asymmetries for tones with increasing and decreasing levels using continuous and global ratings Acta Acust Acust *93*, 623-631.

Sutton, R.S., and Barto, A.G. (1998). Reinforcement learning : an introduction (Cambridge, Mass.: MIT Press).

Treue, S. (2003). Visual attention: the where, what, how and why of saliency. Curr Opin Neurobiol *13*, 428-432.

VanRullen, R. (2003). Visual saliency and spike timing in the ventral visual pathway. J Physiol Paris *97*, 365-377.

Wang, J., Qin, L., Chimoto, S., Tazunoki, S., and Sato, Y. (2014). Response characteristics of primary auditory cortex neurons underlying perceptual asymmetry of ramped and damped sounds. Neuroscience *256*, 309-321.

Womelsdorf, T., and Fries, P. (2007). The role of neuronal synchronization in selective attention. Curr Opin Neurobiol *17*, 154-160.

Zhang, S., Xu, M., Kamigaki, T., Hoang Do, J.P., Chang, W.C., Jenvay, S., Miyamichi, K., Luo, L., and Dan, Y. (2014). Selective attention. Long-range and local circuits for top-down modulation of visual cortex processing. Science *345*, 660-665.

Znamenskiy, P., and Zador, A.M. (2013). Corticostriatal neurons in auditory cortex drive decisions during auditory discrimination. Nature *497*, 482-485.

Figures and legends

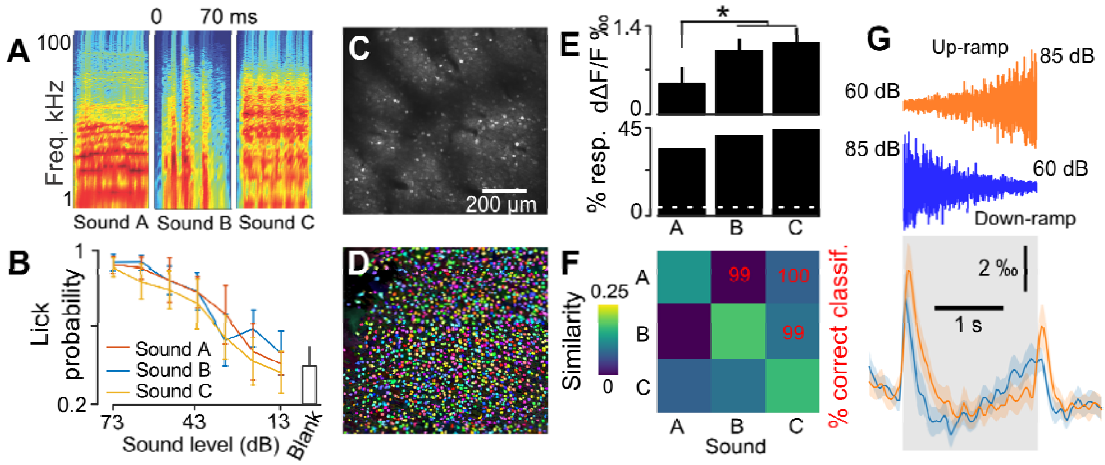

Figure 1: Sounds of equal SPL intensity and detectability recruit different amounts of neural activity in the auditory cortex.

A. Spectrograms of three 70 ms long complex sounds. **B.** Mean response probability for 6 mice trained to detect sounds A, B and C at 73 dB to get a reward and probed with lower sound intensities. While the effect of intensity was significant, there was no effect of sound identity (Friedman test, $p_{intensity} = 2.3\times10^{-9}$, $p_{sound} = 0.43$, n=6 mice). **C.** Example field of view illustrating GCAMP6s labeling of L2/3 auditory cortex neurons. **D.** Result of automated cell segmentation run on the data acquired in the example shown in C. **E.** Mean deconvolved signal (average over stimulus duration) and fraction of significantly responding neurons (Wilcoxon signed test, $p<0.05$) to sounds A, B and C. Mean calcium responses to sound A ($0.05 \pm 0.03\%$ $\Delta F/F.s^{-1}$) were significantly smaller than to B ($0.10 \pm 0.02\%$ $\Delta F/F.s^{-1}$) and C ($0.12 \pm 0.02\%$ $\Delta F/F.s^{-1}$; sign test, $p = 0.0094$ and $p = 0.029$, n = 26 sessions, 15511 neurons in 6 mice). The fractions of responding neurons (35, 41 and 45 %) were all significantly different from each other (χ^2 test, $p = 10^{-12}$, 10^{-27} and 3×10^{-72}, n = 15511). **F.** Population response reliability (diagonal) and similarity (off-diagonal) matrix for sounds A, B and C. The

pair-wise discriminability value, computed with a linear classifier is indicated in red. **G.** Mean deconvolved calcium signals for 6757 auditory cortex neurons in 12 awake mice during 29 calcium imaging sessions for 2 s long white noise sounds modulated in intensity between 60dB to 85dB upwards and downwards.

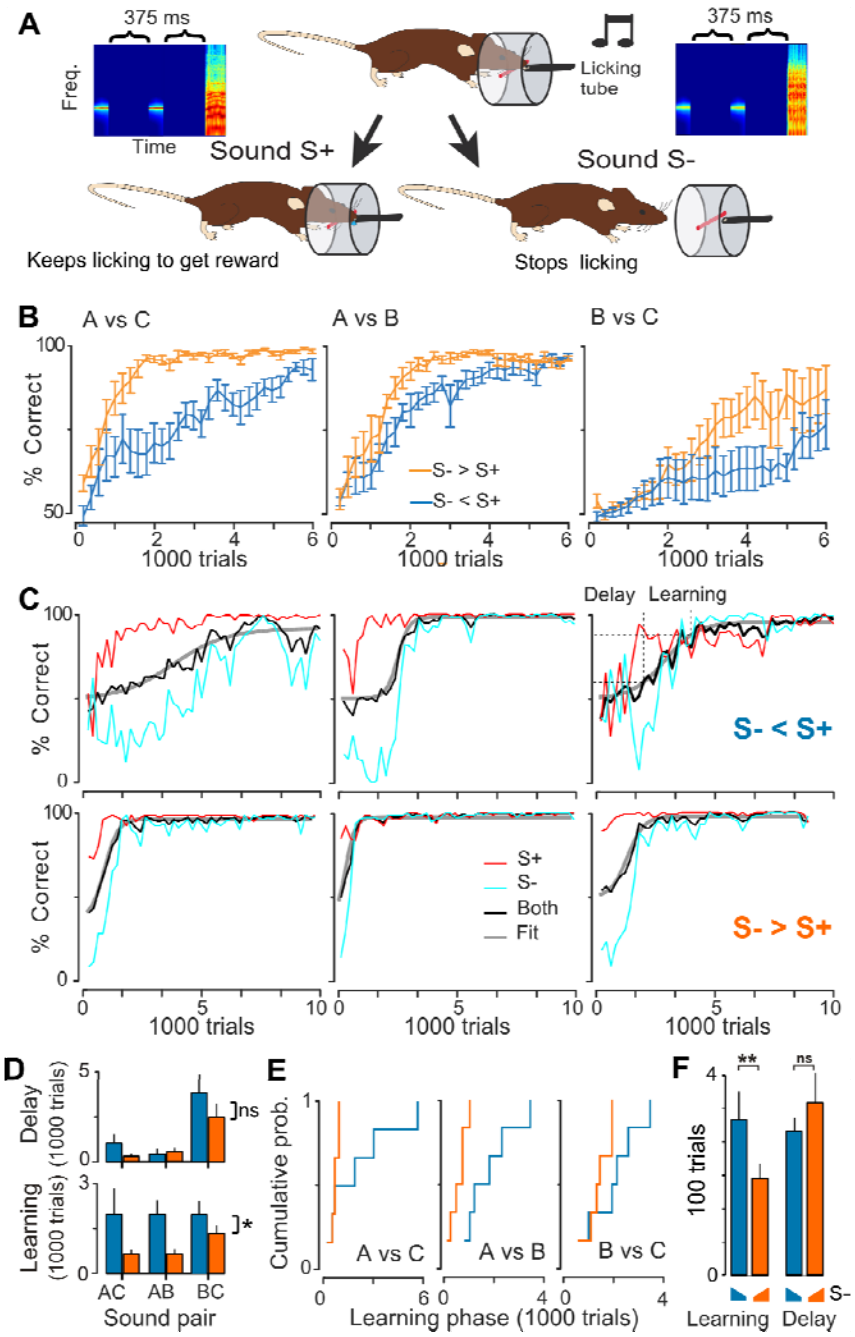

Figure 2: Learning speed is lower and more variable when S- recruits less cortical activity than S+.

A. Schematics describing the auditory Go/NoGo discrimination task. **B.** Mean learning curves for different groups of mice (n = 6 for each curve) discriminating between sounds A and C (left), A and B (middle), B and C (right). The group-curves in orange indicate when the S-

sound has higher levels of cortical recruitment than the S+ sound, while the blue curve stands for the opposite stimulus assignment. **C.** Individual learning curves for 6 mice discriminating sounds A and C. Performance for S+ (red), S- (light blue) and both (black) sounds are displayed. Mice from the top row have sound A as the S- stimulus while mice from the bottom row have sound C as the S- stimulus. Typical learning curves display a delay and learning phase as shown in the top-right graph. **D.** Mean +/- standard error for the delay and learning phase for the three discriminated sound pairs and for groups in which S- recruits more (orange) or less (blue) activity than S+. A significant effect of cortical recruitment for S- vs S+ is found for the learning phase (Friedman test, p=0.016, n=6 mice per condition) but not for the delay phase (Friedman test, p=0.45, n=6 mice per condition). **E.** Cumulative distributions of learning phase durations for sound pairs A-C, A-B and B-C. **F.** Mean +/- standard error for the delay and learning phases for the discrimination of up- and down-ramps (white noise, 2s, range 60 / 85dB) for two groups of mice having the up- (orange) or down- (blue) ramp as S-. A difference is observed for the learning phase (Kolmogorov-Smirnov test, p=0.0046, n=12 mice per condition) but not for the delay phase (Kolmogorov-Smirnov test, p=0.78, n=12 mice per condition).

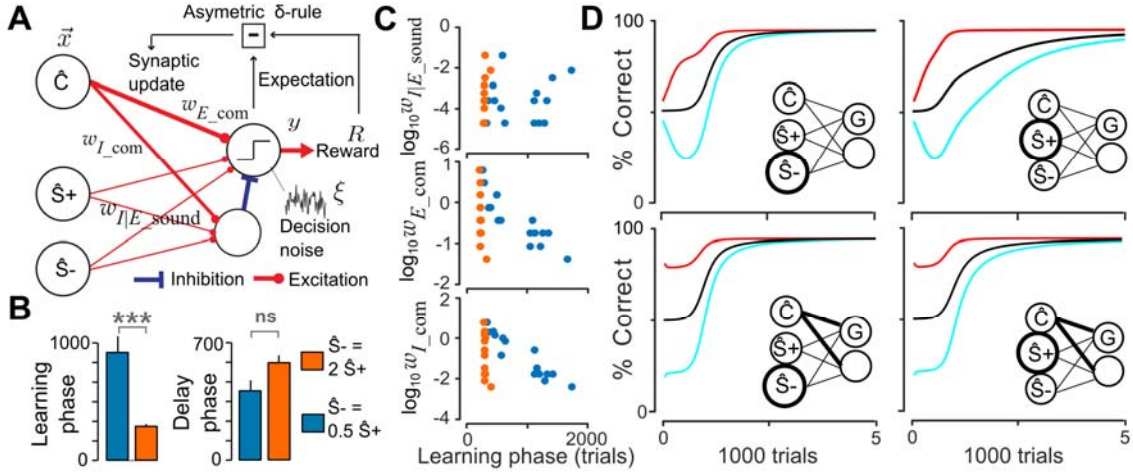

Figure 3: A multiplicative reinforcement learning model explains salience-related variations in learning speed.

A. Schematics describing the auditory Go/NoGo discrimination model. **B.** Mean +/- standard errors for the learning and delay phases obtained with the model initialized with 15 initial conditions obtained by fitting learning curves obtained in a previous study. However, the model now includes a two-fold difference in cortical recruitment between the two stimuli. A longer learning phase (Kolmogorov-Smirnov test, $p=8\times10^{-7}$, n=15 initial conditions per group) but not delay phase (Kolmogorov-Smirnov test, p=0.13, n=15 initial conditions per group) is observed when S- recruits less activity (blue) as compared to when S- recruits more activity than S+ (orange). **C.** Learning phase duration when plotted against the value of the modeled initial synaptic weights. (Same simulations and color code as in B). **D.** Simulated learning curves for four different starting conditions: Strong (bottom row) vs weak (top row) initial connection from unspecific cues and more (left) vs less (right) activity recruited by the S- stimulus.

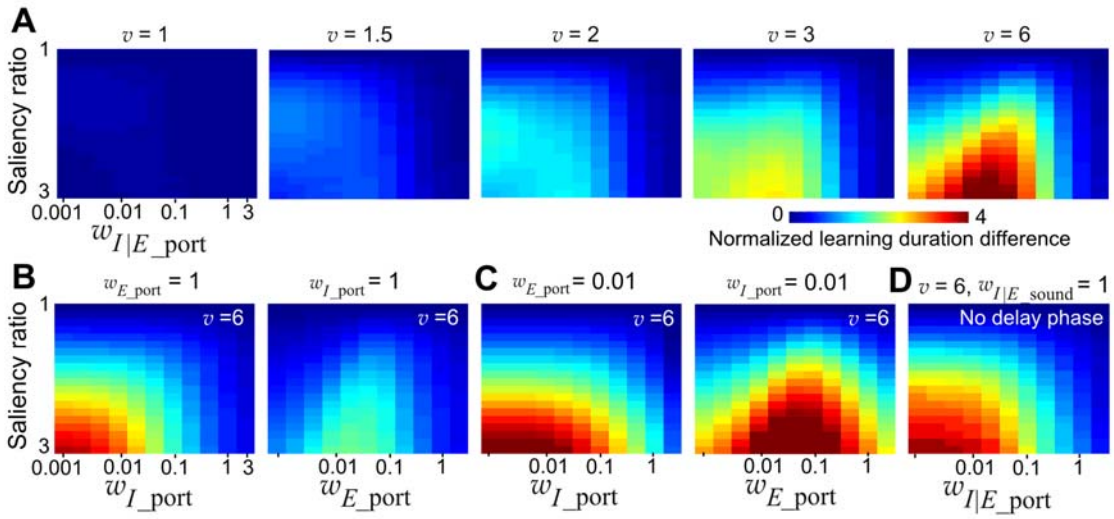

Figure 4: The impact of neuronal recruitment on learning speed depends on the initial strengths of synaptic connections in the model

A. Normalized learning phase duration difference (duration S- more salient - duration S- less salient)/(duration S- more salient) is color-coded and plotted against the cortical recruitment ratio of the discriminated stimuli and the strength of the initial inhibitory and excitatory connection to unspecific cues (w_{I_port} and w_{E_port}, same value for both). From left to right the value of the asymmetric learning rate parameter v is increased. Simulation parameters: $w_{I|E_sounds} = 0.03$, $\alpha = 0.01$, $\sigma = 0.6195$. **B.** Same as A, but varying only the initial inhibitory (left) or excitatory connection (right) while the other connection is kept constant at value 1. **C.** Same as B, but with the non-varied initial connection kept at a constant value of 0.01. **D.** Same as A, but now with initial connectivity from sound stimuli $w_{I_sounds} = 1$ such that there is no delay phase.

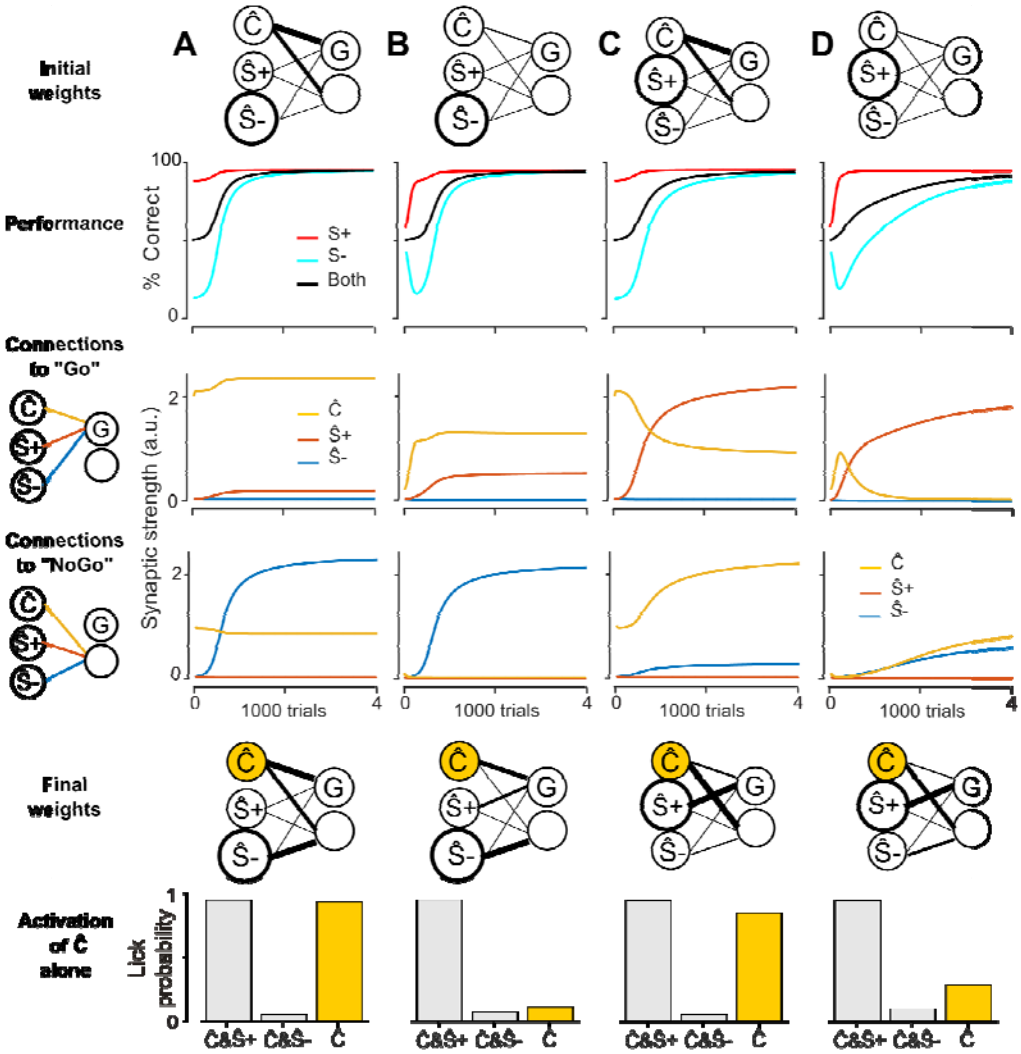

Figure 5: Selection of learning strategy depends on neuronal recruitment in the model.

A. (top) Sketch of the initial synaptic weights and simulated model performance for S+ (red), S- (light blue) and both stimuli (black). (middle)Values of the connections to the excitatory (G) and inhibitory decision populations as indicated by the schematics on the left-hand-side. Yellow: connections from the "common" Ĉ population. Red: connections from the Ŝ+ population. Blue: Connections from the Ŝ- population. (bottom) Sketch of the connectivity pattern after and response probability after learning for the S+ (co-activation of Ŝ+ & Ĉ) and S- (co-activation of Ŝ- and Ĉ) reinforced stimuli as well as for the common stimulus component alone (Ĉ, yellow). Simulation parameters: $\vec{x} = [1\ ;\ 0.66\ ;\ 1.33]$, $\alpha = 0.01$, $\sigma = 0.6195$, $v = 6$, $w_{I_port} = 1$, $w_{E_port} = 2$ and $w_{I|E_sound} = 0.01$. **B.** Same as A, but with $\vec{x} = [1\ ;$

1.33 ; 0.66]. **C.** Same as A, but with w_{I_port} = 0.1, w_{E_port} = 0.2. **D.** Same as B, but with w_{I_port} = 0.1, w_{E_port} = 0.2.

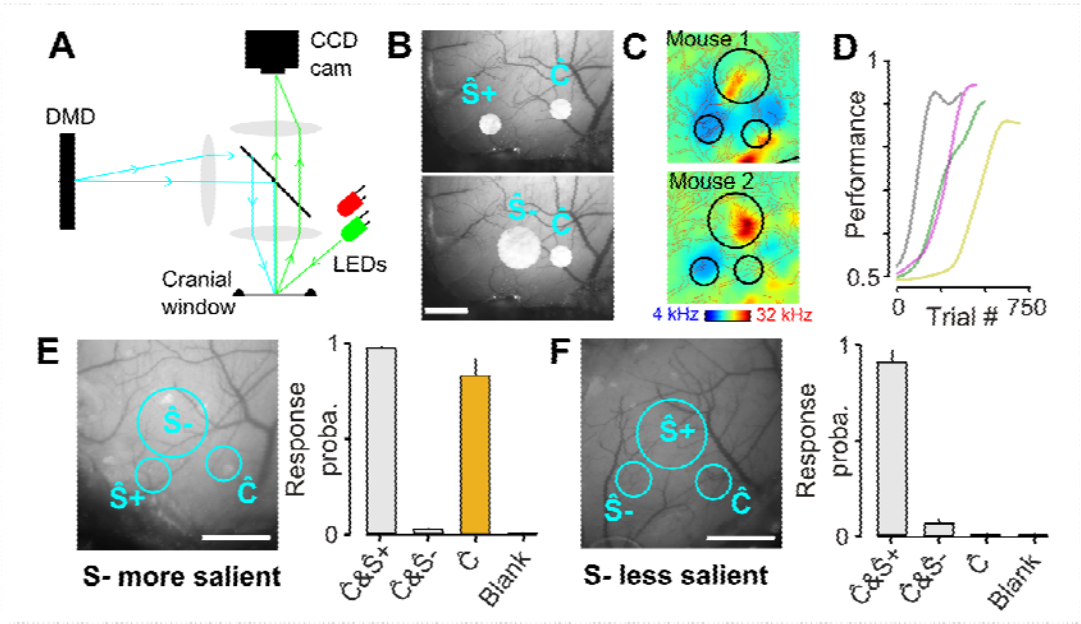

Figure 6: Discrimination training of multi-spot optogenetic patterns reveals a choice of learning strategy depending on the level of cortical recruitment

A. Schematics of the optical setup to project arbitrary 2D light patterns onto the surface of auditory cortex. Blue light patterns from a Digital Micromirror Device (DMD) are collimated and deflected through the objective lens by a beam splitter. The surface of the cranial window can be simultaneously imaged by a CCD camera using external LEDs (green for blood vessel, red for intrinsic imaging). **B.** Examples of light patterns used for the discrimination task. Scale bars: 600 µm. **C.** Tonotopic maps obtained with intrinsic imaging for two mice, and localization of the three optogenetic stimulation spots in the same high, low and mid frequency fields. **D.** Four example learning curves for the optogenetic discrimination task. Grey: smaller Ŝ - ; color: larger Ŝ -. **E.** (left) Light pattern in the task in which the S- stimulus has higher level of cortical recruitment (Ŝ- >Ŝ+). Ĉ = common component of S+ and S- stimuli. (right) Response probability on the two learnt target stimuli (Ĉ & Ŝ + vs Ĉ & Ŝ -) and

on presentation of the common part of the stimulus alone in 15 catch trials (yellow). Mean ± SEM, n=4 mice. **F.** Same as E, but with Ŝ + larger than Ŝ -. Mean ± SEM, n=4 mice. Scale bars: 600 μm.

STAR methods.

CONTACT FOR REAGENT AND RESOURCE SHARING

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Brice Bathellier (brice.bathellier@unic.cnrs-gif.fr).

EXPERIMENTAL MODEL AND SUBJECT DETAILS

All mice used for imaging and behavior were 8 to 16 weeks old C57Bl6J and GAD2-Cre (Jax #010802) x RCL-TdT (Jax #007909) mice. Mice used for optogenetics were 8 to 16 weeks old males and female obtained by crossing homozygous Emx1$^{\text{IRES-cre}}$ (Jax #005628) mice with Ai27 (Jax # 012567) mice to obtain expression of Td-Tomato-tagged channelrhodopsin (ChR2) in excitatory neurons of the cortex. All animal were group housed. All procedures were approved by the Austrian laboratory animal law guidelines (Approval #: M58 / 02182 / 2007 /11; M58 / 02063 / 2008 /8) and the French Ethical Committee (authorization 00275.01).

METHOD DETAILS

Two-photon calcium imaging in awake mice. At least three weeks before imaging, mice were anaesthetized under ketamine medetomidine. The right *masseter* was removed and a large craniotomy (~5 mm diameter) was performed above the auditory cortex. We then performed three to five injections of 200nL (35-40nL/min), rAAV1.Syn.GCaMP6s.WPRE virus obtained from Vector Core (Philadelphia, PA) and diluted 10 times. The craniotomy was sealed with a glass window and a metal post was implanted using cyanolite glue and dental cement. At least three days before imaging, mice were trained to stand still, head-fixed under the microscope for 20 to 60 min per day. Then mice were imaged one to two hours per day. Imaging was performed using a two-photon microscope (Femtonics, Budapest, Hungary)

equipped with an 8kHz resonant scanner combined with a pulsed laser (MaiTai-DS, SpectraPhysics, Santa Clara, CA) tuned at 920 nm or 900nm depending on the experiments. Images were acquired at 31.5Hz. All sounds were delivered at 192 kHz with a NI-PCI-6221 card (National Instrument) driven by Elphy (G. Sadoc, UNIC, France) through an amplifier and high frequency loudspeakers (SA1 and MF1-S, Tucker-Davis Technologies, Alachua, FL). Sounds were calibrated in intensity at the location of the mouse ear using a probe microphone (Bruel&Kjaer). In a first experiment, we played three 70 ms complex sounds at 73dB SPL preceded by two 50 ms 4kHz pure tones (inter-tone interval: 375 ms) sounds as in the behavioral task. The three sounds were played in a random order and repeated 30 times. In a second experiment, we played white noise sounds ramping -up or -down in intensity between 60dB and 85dB SPL during 2s. The ramps were repeated 20 times.

Data analysis. Data analysis was performed using Matlab and Python scripts. Motion artifacts were first corrected frame by frame, using a rigid body registration algorithm. Regions Of Interest were selected using a semi-automated hierarchical clustering algorithm based on pixel covariance over time as described in (Roland et al., 2017)(see detailed method below). Neuropil contamination was subtracted (Kerlin et al., 2010) by apply the following equation: F_{true} (t) = $F_{measured}$ (t) – 0.7 $F_{neuropil}$ (t), then the change in fluorescence ($\Delta F/F_0$) was calculated as $(F-F_0) / F_0$, where F_0 is estimated as the minimum of the low-pass filtered fluorescence over ~40 s time windows period. To estimate the time course of firing rate, the calcium signal was temporally deconvolved using the following formula: r(t) = f'(t) + f(t) / τ in which f'(t) is the first derivative of f(t) and τ the decay constant set to 2 seconds for GCaMP6s. For the complex sounds, the population response was computed as the mean deconvolved signal across all neurons from sound onset to 500ms after sound offset. For the ramps, because the behavioural discrimination response typically occurs within few hundreds

of milliseconds after the ramp onset (Deneux et al., 2016), the mean population response was evaluated from 0 to 500ms after sound onset.

Patterned optogenetics and intrinsic imaging. To flexibly activate different activity patterns in the mouse auditory cortex, we used a computer driven (VGA input) video projector (DLP LightCrafter, Texas Instruments) which includes a strong blue LED light source (460 nm) and from which we have removed the objective. To project a two-dimensional image onto the auditory cortex surface (**Figure 6A, B**), the image of the micromirror chip is collimated through a 150 mm cylindrical lens (Thorlabs, diameter: 2 inches) and focused through a 50 mm objective (NIKKOR, Nikon). Imaging of the cortex at the focal plane is obtained by side illumination with a green (525 nm, blood vessels) or far red (780 nm, intrinsic imaging) LED. The light collected by the objective passes through a dichroic beamsplitter (long pass, >640nm, FF640-FDi01, Semrock) and is collected by a CCD camera (GC651MP, Smartek Vision) equipped with a 50 mm objective (Fujinon, HF50HA-1B, Fujifilm). Note that the image projected to the cortical surface corresponds to a narrow cone of light extending below the surface and potentially activating ChR2 expressing neurons throughout the cortical depth. Intrinsic imaging was performed in isoflurane anesthetized mice (1.1% delivered through SomnoSuite, Kent Scientific). To compute intrinsic signal maps we divided the red light image of the cortical surface after the onset of a stimulation (average over 2 s) with 2s long pure tones (4, 8, 16 and 32kHz) by the mean image immediately before stimulus onset

Go/NoGo discrimination behavior. Mice were water-deprived and trained daily for 200 to 300 trials. Mice first performed 4 habituation sessions to learn to obtain a water rewards (~5 µl) by licking on a spout over a threshold after the positive stimulus S+. After habituation, the

fraction of collected rewards was about 80%. The learning protocol then started in which mice also received a non-rewarded, negative sound S- for which they had to decreasing licking below threshold to avoid an 8s time-out. For the freely moving complex sound discrimination, S+ and S- sounds consisted of two 4 kHz pips (50 ms) followed by one of the three 70 ms complex click shown in **Figure 1A**. The interval between the offset and onset of the pips and click was 375 ms. Licking was assessed 0.58 sec after the specific sound cue in a 1 sec long window by an infrared beam at the spout. For the intensity ramp discrimination, licking was assessed in a 1.5s window after sound offset. In both cases, licking was considered above threshold if the infrared beam in front of the licking tube was broken during 75% of the measurement time-window. Positive and negative sounds were played in a pseudorandom order with the constraint that exactly 4 positive and 4 negative sounds must be played every 8 trials. For learning curves, performance was measured as the fraction of correct positive and correct negative trials over bins of 100 trials. For the optogenetically-driven, head-fixed discrimination task, the S+ and S- stimuli were each composed of two disks of blue light (465 nm) flashing at 20 Hz for 1 s. One of the two disks (noted Ĉ) was common to S+ and S- stimuli, the other disk was condition-specific. The three disk locations were chosen in similar tonotopic locations across mice based on intrinsic imaging maps. They were precisely re-positioned for every training session using an automated registration procedure based on blood-vessel patterns. A strong masking light was used to prevent the animal from using visual cues in the task. The common disk was 360 µm in diameter. In one set of mice, the disk specific to S- was 720 µm in diameter, while the S+ specific disk was 360 µm in diameter. In the other set of mice, sizes of the specific disks were swapped. Head-fixed mice performed 200 to 300 trials per day with an inter-trial interval randomized between 3 s and 7 s. Individual licks were detected through an electric circuit connecting the mouse and the lick tube. Then, each trial was started only if the mouse was not spontaneously licking for at least

3 s (in3 addition to the inter-trial interval). Mice were first trained to respond to the S+ stimulus by producing at least 3 to 5 licks (depending on the mouse) to get the 5 µL water reward. When the mouse could collect more than 80% of the rewards, the S- stimulus was introduced. Licking above threshold after S- was punished with a 7 s timeout.

Reinforcement learning model. The model has been described extensively in a previous publication (Bathellier et al., 2013). In short, it is composed of three sensory units (Ŝ+, Ŝ- and Ĉ, representing populations of neurons) whose activity described by a three-dimensional vector $\vec{x}$ and which are connected to a simple decision circuit (**Figure 3A**). Cortical recruitment is modeled by changing the firing value of the sound units. When the S- stimulus recruits less activity than the S+ stimulus, the input vectors are: $\vec{x}_{S+}$ = [1 0 2] or $\vec{x}_{S-}$ = [1 1 0]. When S- recruits more activity than S+, the input vectors are: $\vec{x}_{S+}$ = [1 0 1] or $\vec{x}_{S-}$ = [1 2 0]. In **Figure 4**, the recruitment ratio is equal to the larger of the two firing rate values.

The decision circuit is composed of all-or-none response unit (y = 0 or 1) which linearly sums the three sensory inputs (representing synaptic populations) under the form of three direct excitatory connections and of a graded feed-forward inhibition from a virtual inhibitory unit in fact equivalent to three direct inhibitory connections. The output of model is described by a single equation for the decision unit:

$$y = \theta(\vec{w}_E.\vec{x} - \vec{w}_I.\vec{x} - \xi)$$

in which θ is the Heaviside step function. $\vec{w}_E$ and $\vec{w}_I$ are three-dimensional positive vectors describing the excitatory synaptic weights from the sensory units to the decision and inhibitory units respectively. The variable ξ is a Gaussian random noise process of unit variance which models the stochasticity of behavioral choices.

Based on the action outcome (R = 1 for a reward, R = -1 for no reward), the learning rule for the synaptic weights is implemented as:

$$\delta\vec{w}_E = \alpha\vec{w}_E \circ f[R - \sigma(\vec{w}_E - \vec{w}_I).\vec{x}]y\vec{x}$$

$$\delta\vec{w}_I = -\alpha\vec{w}_I \circ f[R - \sigma(\vec{w}_E - \vec{w}_I).\vec{x}]y\vec{x}$$

in which $\circ$ is the Hadamard (element-wise) product implementing the multiplicative rule, $y\vec{x}$ is a Hebbian term, α is the learning rate and σ is a parameter related to the noisiness of the model and setting its asymptotic performance. To account for the faster improvement of performance for rewarded as compared to non-rewarded trials, positive expectation errors are more strongly weighted than negative ones, thanks to the asymmetric function $f[u] = u$ if $u \leq 0$ and $f[u] = \nu u$ if $u > 0$. The parameter ν is typically larger than 1, consistent with the activity of basal ganglia dopaminergic neurons in mice (Cohen et al., 2012) and monkeys (Schultz et al., 1997) coding for reward expectation error.

As described above, the equations of the model are stochastic due to the Gaussian random noise process ξ. To compute the response probability estimates plotted throughout the study, we used a previously established probability equation (Bathellier et al., 2013), valid for learning dynamics much slower than fluctuations (ergodic approximation). The probability to make a lick response given the input vector $\vec{x}_{S+}$ or $\vec{x}_{S-}$ is:

$$p_{S+|S-} = p(y = 1 | \vec{x} = \vec{x}_{S+|S-}) = \frac{1}{2}\left(1 + \mathrm{erf}\left(\frac{\Delta\vec{w}.\vec{x}_{S+|S-}}{\sqrt{2}}\right)\right)$$

where $\Delta\vec{w} = \vec{w}_E - \vec{w}_I$ represents now the average observed values of difference between the excitatory and inhibitory connections. In addition, the plasticity equations become:

$$\delta\vec{w}_E = \frac{\alpha}{2}\vec{w}_E \circ \left(f[1 - \sigma\Delta\vec{w}.\vec{x}_{S+}]p_{S+}\vec{x}_{S+} + f[-1 - \sigma\Delta\vec{w}.\vec{x}_{S-}]p_{S-}\vec{x}_{S-}\right)$$

$$\delta\vec{w}_I = -\frac{\alpha}{2}\vec{w}_I \cdot \left(f[1-\sigma\Delta\vec{w}.\vec{x}_{S+}]p_{S+}\vec{x}_{S+} + f[-1-\sigma\Delta\vec{w}.\vec{x}_{S-}]p_{S-}\vec{x}_{S-} \right)$$

QUANTIFICATION AND STATISTICAL ANALYSIS

All quantifications are given as mean ± standard error (SEM). To statistically assess the differences between paired measurements (e.g. activity for two different sounds elicited in the same neuronal populations) we used the non-parametric Wilcoxon signed rank test. To compare two sets of measurements (e.g. delay and learning phase duration for two groups of mice) we used the non-parametric Wilcoxon rank sum test. Assessment of the differences in the fraction of responsive neurons for different sounds was done with the χ^2 test which evaluates differences in the distributions of two binary variables.

DATA AND SOFTWARE AVAILABILITY

Datasets, analysis software and codes for running the simulations of our model are available upon request to Brice Bathellier (brice.bathellier@unic.cnrs-gif.fr).

KEY RESOURCES TABLE

Separate file.

Supplemental items

Supplemental Figure S1: **Cortical recruitment differences are robust across mice and experiments.** related to Figure 1

Supplemental Experimental Procedures: analytical arguments in support of the simulations presented in the paper.