

Suppression of cortical representation through backward conditioning

Shaowen Bao, Vincent T. Chan, Li I. Zhang, and Michael M. Merzenich*

Keck Center for Integrative Neuroscience, University of California, San Francisco, CA 94143

Contributed by Michael M. Merzenich, December 10, 2002

Temporal stimulus reinforcement sequences have been shown to determine the directions of synaptic plasticity and behavioral learning. Here, we examined whether they also control the direction of cortical reorganization. Pairing ventral tegmental area stimulation with a sound in a backward conditioning paradigm specifically reduced representations of the paired sound in the primary auditory cortex (AI). This temporal sequence-dependent bidirectional cortical plasticity modulated by dopamine release hypothetically serves to prevent the over-representation of frequently occurring stimuli resulting from their random pairing with unrelated rewards.

Cortical stimulus feature representations can be reorganized by our experiences throughout life. In the auditory system, for example, frequency discrimination training results in a selective expansion of the cortical area representing the training frequencies (1). Simple auditory classical conditioning and instrumental conditioning can also modify the spectral receptive fields of cortical neurons, enhancing and suppressing responses evoked by the paired sound (2–9). Although response suppression effects have been puzzling, in general, these cortical plasticity effects are believed to underlie improved sensory discrimination, to enlarge and temporally organize the responding population, and to enhance the salience of conditioned stimuli (10–12).

The activation of brain neuromodulator systems appears to be required for long-lasting plasticity in the adult cerebral cortex (7, 13–18). Brain cholinergic and dopaminergic systems are activated in the cortex during learning (19–22) and have been demonstrated to play important roles in cortical plasticity (7, 15, 23–29). Cholinergic basal forebrain lesions or cortical cholinergic receptor blockade prevents cortical reorganization (13–18). Systematically pairing the microstimulation of cholinergic basal forebrain or dopaminergic ventral tegmentum neurons with sensory stimulation induces distinctly different forms of cortical plasticity that parallel different aspects of change induced by behavioral training (1–4, 15, 23–28). The specific form of the cortical plasticity is likely determined by a number of factors, including the spectral, spatial, and temporal characteristics of the sensory input and the temporal sequence of the sensory and neuromodulator activation (15, 24–26).

Studies of experience-dependent plasticity have demonstrated enhancement of cortical representations of behaviorally important stimuli, in which neuromodulators are involved presumably through a positive modulation of long-term potentiation (30–33), a putative cellular mechanism underlying plasticity (34, 35). Under certain circumstances, neuromodulators also enhance long-term depression (36, 37), suggesting that they may also mediate reduction of cortical representation.

The temporal relationships of stimulus and reinforcement appear to be important for determining the directions of brain plasticity. On the synaptic level, synaptic transmission is potentiated if presynaptic activity consistently precedes postsynaptic action potentials and is depressed if presynaptic activity trails postsynaptic action potentials (38, 39). On the cortical level, directions of cortical plasticity are also determined by timing of the stimulus and cortical activity (40, 41). On the behavioral

level, stimulus-reinforcement contingency determines the directions of learning. Although forward conditioning [reward follows the conditioned stimulus (CS)] results in the learning of conditioned responses (CRs), backward conditioning (rewards precede the CS, a procedure of negative contingency) impairs subsequent learning (42). Given that forward conditioning generally increases cortical CS representations (2, 3, 15, 24, 27), we here examined whether backward pairing of ventral tegmental stimulation and auditory stimuli systemically reduces the cortical representations of paired stimuli.

Methods

Preparation. Platinum bipolar stimulating electrodes were stereotaxically implanted within the right ventral tegmental area (VTA) (4.5 mm posterior, 0.7 mm lateral, and 8.5 mm ventral to Bregma) in barbiturate-anesthetized female rats (300 g) by using techniques approved under University of California, San Francisco, Animal Care Facility protocols. After a 2-week recovery period, rats were placed in an operant conditioning chamber and were allowed to bar press for brief VTA microstimulation (10 biphasic pulses of 0.1-ms duration at 100 Hz). The minimal current levels that reinforced consistent bar presses at least once every 2 sec were determined as the electrical stimulus threshold (100–200 μ A). Subsequent pairing of auditory stimuli with VTA stimulation took place in a sound-attenuation chamber. One group of six rats was presented with paired VTA electrical microstimulation (10 biphasic pulses of 0.1-ms duration at 100 Hz) and a narrow band pulsed noise (six 25-ms noise pulses with 5-ms on/off ramp delivered at a rate of 10 pips/sec, 55 dB SPL peak intensity, in the frequency band of 7.3–11.1 kHz, started 200 or 500 ms after the offset of the VTA stimulation). Each daily session consisted of 360 pairing trials. The intervals between successive pairing trials were pseudorandom in the range from 12 to 28 sec. Twenty sessions were given in a 4-week period. A second group of three animals underwent the same pairing procedure as the first, except that the noise was replaced with 9-kHz tone. To control for stimulus-induced changes in the auditory cortex, four animals were presented with a 9-kHz pulsed tones and four animals with pulsed noise alone for the same number of sessions.

Electrophysiology. Twenty-four hours after the last stimulation session, experimental and auditory control animals were anesthetized with sodium pentobarbital, the right auditory cortex surgically exposed, and neuronal responses recorded with parylene-coated tungsten microelectrodes (FHC, Bowdoinham, ME, <1- μ m tip diameter, 1–2 M Ω at 1 kHz). Six naïve animals underwent similar procedures. Recording sites were chosen to evenly sample from the auditory cortical zone while avoiding blood vessels. At every site, the recording microelectrode was lowered orthogonal to the surface 470–550 μ m in depth (layers 4/5), where vigorous driven responses were recorded. Only one recording was made per penetration. The evoked spikes of a

Abbreviations: VTA, ventral tegmental area; BF, best frequency; CR, conditioned response; CS, conditioned stimulus.

*To whom correspondence should be addressed. E-mail: merz@phy.ucsf.edu.

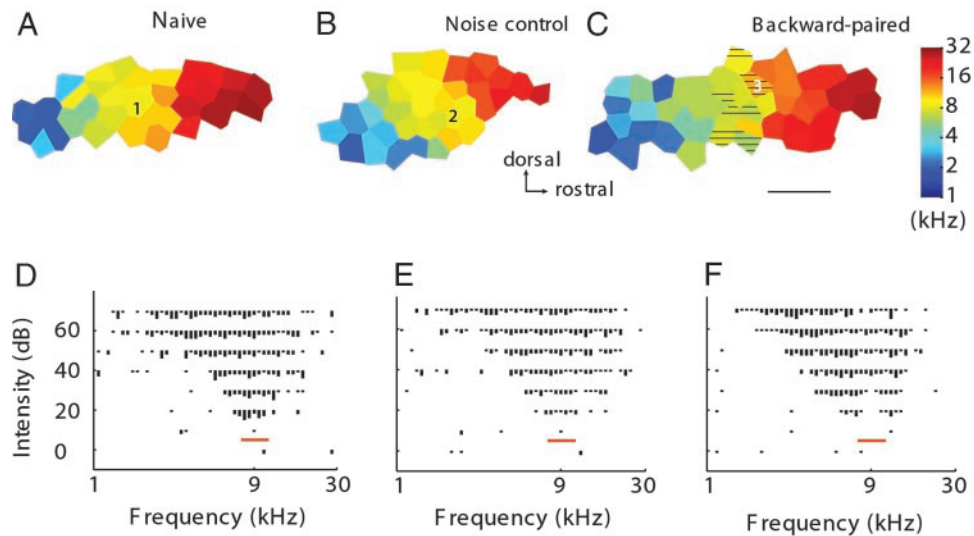


Fig. 1. Reduced cortical responses to tones in the backward-paired frequency range. (A–C) Cortical maps from a naïve, a noise control, and a noise–VTA backward-paired animal. Hatched areas exhibited a notch in their receptive fields at the paired noise frequency band. (D–F) Receptive fields recorded from sites marked 1–3 in the representative maps. The receptive field recorded from the backward-paired animal had double peaks caused by a notch at the paired noise frequency band (marked with a red bar).

neuron or a small cluster of two to five neurons were collected at each site. Spectral-intensity receptive fields were reconstructed in detail by presenting 50 pure-tone frequencies (1–30 kHz, 25-ms duration, 5-ms ramps) at each of eight sound intensities to the contralateral ear at a rate of two stimuli per sec by using a calibrated sound delivering system. The tuning curve characterization was made by using a “blind” procedure, in which all receptive field files were pooled and presented in a computer-generated random sequence. The sequence was stored in a file for later “decoding” of the results. An objectively recorded best frequency (BF) was defined as the frequency that evoked a neuronal response at the lowest stimulus intensity. To generate BF maps, points on the cortex were assigned the BFs of the nearest recording site through Voronoi tessellation. The boundaries of the map were functionally determined by using sites that did not have a well-defined frequency-intensity receptive field. The response latency was defined as the time from stimulus onset to the earliest response ($4 \times \text{SD}$ above baseline activity) for five frequencies that were nearest the BF at a sound level (30 or 70 dB SPL). Response magnitude was defined as the average number of spikes per tone for five frequencies nearest the BF at a sound level (30 or 70 dB SPL). Response latencies and magnitudes were analyzed only for sites in the tonotopic BF maps (i.e., sites with well-defined receptive fields). A computer program was used to objectively identify the receptive fields with notches that were at least 20 dB deep, generated by the backward-paired frequency band. The percent of cortical area responding to tones of various frequencies and intensities was calculated by using raw receptive field data of all sites with best frequencies in a range from 5 to 16 kHz.

Unless otherwise specified in the text, statistical significance was assessed by using a two-tail *t* test. Data are presented as mean \pm SEM. On completion of the experiment, electrolytic lesions were made through the stimulating electrodes, and the placement of the stimulating electrodes in the VTA was histologically verified.

Results

A group of rats ($n = 6$) were presented with paired VTA electrical microstimulation and pulsed band-limited noise (7.3–11.1 kHz) in a backward conditioning sequence (VTA stimula-

tion preceding the auditory stimulus) for 20 days. Control animals were naïve or received only noise stimulation ($n = 4$). Animals receiving only VTA stimulation have been previously shown to exhibit no significant auditory cortical reorganization (24). The auditory cortex was mapped in detail 24–48 h after the last conditioning stimulation session. The primary auditory cortex of backward-paired and auditory control animals showed largely normal tonotopic BF organization, and the area of the primary auditory cortex (AI) was comparable to that of naïve animals (Fig. 1 A–C; naïve: 1.24 ± 0.09 , $n = 6$; noise control: 1.20 ± 0.19 , $n = 4$; backward-paired: 1.21 ± 0.28 , $n = 6$, $P > 0.1$). We also quantified neuronal response properties such as response magnitude (number of impulses per stimulus) and latency (see *Methods* for details) at two sound intensity levels (30 or 70 dB SPL). Statistical comparisons were made for all AI sites, sites tuned to $9 \text{ kHz} \pm 0.3$ octave and sites tuned to $9 \text{ kHz} \pm 0.6$ octave. No significant differences were found among the three groups for any of the six analyses (ANOVAs, $P > 0.1$ for all: two intensities and three BF ranges). However, inspection of frequency-intensity receptive fields revealed that those for many recording sites in the backward-paired auditory cortex showed a notch in the range of the paired noise frequency band (Fig. 1F). A computer program was used to objectively define the number of sites with a notch at least 20 dB deep. Although notches were rarely seen in naïve or noise control animals, there were, on average, four sites with a notch in each of the backward-paired animals (naïve: 0.5 ± 0.33 ; noise control: 0.75 ± 0.55 ; backward-paired: 4.3 ± 1.04 , $P < 0.01$). All were located in the AI zone that represented frequencies in the 6–14 kHz, i.e., the paired noise frequency range. These notches manifested reduced cortical responses to the paired noise band.

Reduction of cortical responses in the paired frequency range was also recorded in BF distributions along the tonotopic frequency axis. While BFs were evenly distributed for the noise control group as for the naïve group, the distribution was discontinuous in backward-paired animals (Fig. 2A). There are fewer sites tuned to frequencies close to 9 kHz. The reduction appeared to be limited to a 0.3-octave-wide frequency band. When analyzed with 0.6-octave-wide bins, no significant differences were found in the percentage of cortical area representing any of the eight frequency bands (at 9 kHz: naïve, $23.4 \pm 4.1\%$;

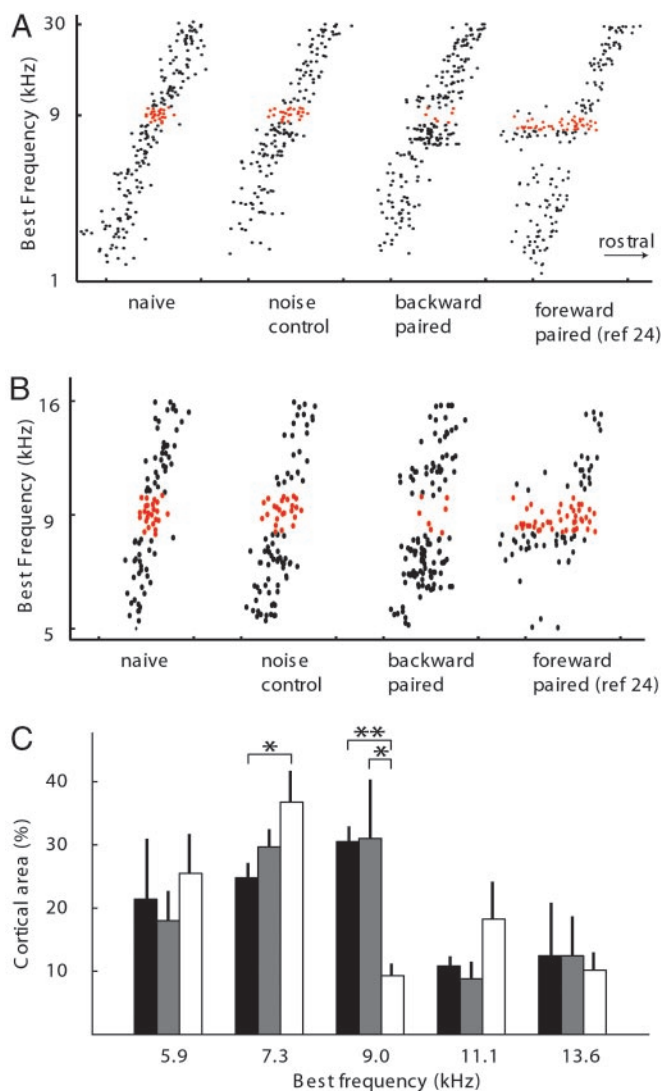


Fig. 2. Reorganization of the auditory cortex by noise-VTA backward pairing. (A) Distribution of BF along the rostrocaudal axis of the auditory cortex. Points with BF in a 0.3-octave range centered at 9 kHz are indicated with red ($n = 6$ for backward-paired groups, $n = 4$ for naïve and noise control; a previously published forward-paired group was also included for comparison, see figure 2 of ref. 24). (B) Similar graph to A on a finer frequency scale (5–16 kHz). (C) Percent of the auditory cortex that was tuned to each frequency band. Only points with BF in the 5- to 16-kHz range are included. Black bar, naïve; gray bar, noise control; white bar, backward-paired; bin size, 0.3 octave ($n = 6$ for naïve and backward-paired groups, $n = 4$ for noise control; *, $P < 0.05$; **, $P < 0.0005$).

noise control, $23.7 \pm 5.7\%$; backward-paired, $16.5 \pm 4.1\%$, $P > 0.1$). On the other hand, for a 0.3-octave bin size, there was a significant reduction in the percentage of cortical area representing a band centered at 9-kHz in the backward-paired animals compared with naïve (Fig. 2B, $P < 0.0005$) and noise control animals ($P < 0.05$). In addition, areas representing the adjacent frequencies (0.3 octave centered at 7.3-kHz and 11.1-kHz) appeared to be increased in the backward-paired animals, although statistical significance was reached only for the 7.3-kHz bin ($P < 0.05$). These results do not indicate that only the center half of the 0.6-octave backward-paired noise bandwidth was effective in suppressing cortical representations, because, as noted below, changes in receptive fields are not fully quantified by BF shifts.

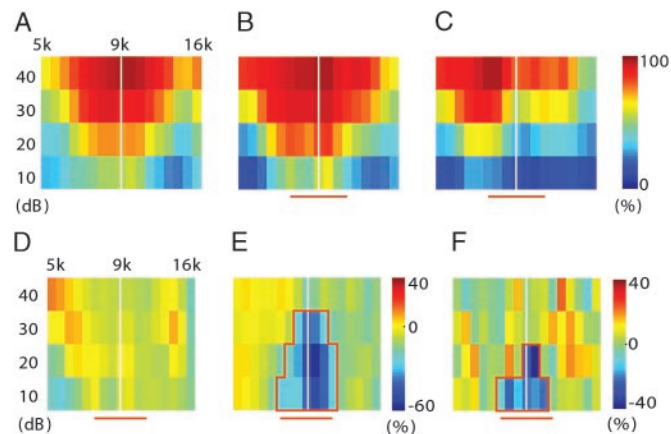


Fig. 3. Noise-VTA backward pairing reduces cortical responses to the paired noise frequency band. (A–C) Percent of cortical area responding to tones of various frequency and intensity calculated for a naïve animal (A), a noise control (B), and a backward-paired animal (C). Only areas with BF in the 5- to 16-kHz range are included in the analysis. (D and E) Change in percent cortical area responding to various tones, calculated by subtracting A from B (D) and B from C (E). Pixels that are 2 SD below average ($P < 0.05$) are enclosed with red lines. (F) Difference in percent cortical area responding to various tones between noise control group ($n = 4$) and backward-paired group ($n = 6$). The red horizontal bars depict the frequency band of the paired noise.

To assess the effects of backward pairing on receptive fields of the overall population of cortical neurons, we calculated the percentage of the cortex responding to frequency-intensity combinations. Only areas with BFs ranging from 6 to 14 kHz were included in this analysis, because neurons tuned to other frequencies were generally not responsive to the paired noise and were unlikely to be affected. Because raw receptive field data instead of experimenter-determined tuning curves were used, and because all sites responsive to the paired noise frequency band were pooled, this analysis is unbiased. Fig. 3A–C show three representative cases. Noise exposure alone did not consistently alter cortical response to the paired frequency band (Fig. 3D). Compared with noise exposure, backward pairing reduced the percentage of cortical area responding to tones in the paired frequency range by as much as 60% (Fig. 3E). Although the reduction of responses was most pronounced in a 0.3- to 0.4-octave bandwidth around 9 kHz, responses in the whole backward-paired frequency band were significantly suppressed (Fig. 3C, $P < 0.05$ for an area of 17 pixels; pixel size, 10 dB \times 0.1 octave). The effect was significant when backward-paired and noise control groups were compared (Fig. 3F).

To test whether the observed reduction of cortical representation was unique to noise, we paired VTA stimulation with a trailing 9 kHz pulsed tone ($n = 3$). An analysis of receptive fields, similar to that in Fig. 3, indicated that significantly less cortical area responded to 9 kHz ($P < 0.05$ for five pixels). The reduction was limited to a bandwidth of 9 kHz \pm 0.15 octave (data not shown).

Discussion

We have previously reported that when VTA microstimulation is preceded by a 4-kHz tone and followed by a 9-kHz tone, the representations of the 4-kHz tone are enhanced and those of the 9-kHz are reduced (24). However, it was unclear whether the reduction was caused by the expansion of the nearby cortical representation of the preceding tone. In the present study, we have demonstrated that backward pairing of VTA stimulation with an auditory stimulus highly selectively suppresses the representation of the paired stimulus. The reduction of cortical representation of a frequency band was accompanied by enhancement of representations of the flanking frequency bands, which appears to be a reversal of the plasticity pattern observed

in forward-paired animals (24). The reduction of cortical responses to backward-paired tones observed in this study is less robust but more frequency-specific than that observed in the two-tone pairing study, suggesting that the reduction of the representation of the tone trailing VTA stimulation observed in the previous study is partly due to the expansion of the representation of the preceding tone.

The suppression of cortical representation induced by backward conditioning is different from simple habituation. Repetitive exposure to an auditory stimulus has been shown to alter the organizations of the auditory cortex and the functions of the auditory system. However, the effects are variable. For example, reduction, no change, and mild enhancement of the response to the exposed tone have all been reported (15, 24, 27, 43, 44). Behaviorally, exposure impairs performance in some tasks (42), but facilitates it in others (45). Furthermore, unlike long-lasting backward pairing-induced suppression, both behavioral and physiological habituations spontaneously recover (43, 44, 46).

The cerebral cortex reorganizes to optimally represent behaviorally important stimuli. The importance of a stimulus is not correlated with its occurring frequency but is often marked by activity of neuromodulator systems such as basal forebrain cholinergic and ventral tegmental dopaminergic systems. Repeated pairing of a stimulus with cholinergic or dopaminergic activity results in enhanced representations of the stimulus. However, a mechanism that increases cortical representations based only on temporal coincidence would cause overrepresentation of frequently occurring stimuli because of the high probability of random pairing between such stimuli and unrelated rewards. By contrast, a contingency-based mechanism, which increases cortical representations on positive contingency and decreases them on negative contingency, would selectively enhance representations of behaviorally important stimuli. An irrelevant stimulus, regardless of its occurring frequency, would have no contingency with unrelated rewards, and its representations would not be modulated.

In some reinforcement learning, animals have to make specific motor responses when a CS is presented to obtain a reward (e.g., pressing a bar after hearing a tone). The ventral tegmental dopa-

minergic neurons initially respond to the reward. As the animal learns, the VTA dopaminergic neurons start responding to the CS before the CR and the reward (47). Our results seem to suggest that the VTA response to the CS could adversely impact the cortical representations of the ensuing CR. Learning motor responses is generally initially associated with enhanced neural responses and enlarged movement representations in the motor cortex (48–50). As the behavior comes to be performed “automatically,” i.e., without requisite attention, the initial changes may fade in the cortex to drive it in the direction of its initial pretraining representational state. If these modulatory effects of dopamine release contribute to active “unlearning,” the transition in timing of dopamine release between dopamine–CR (backward) and CR–dopamine (forward) sequences must be directly linked to processes governing CS–CR automaticity. Further studies are needed to confirm whether this is the case.

It is also important to note that successful performance of the CR is a prerequisite for the CS to predict the reward and to activate the VTA. Failure to perform CRs will result in, (i) disruption of the CS–reward prediction; (ii) activation of VTA dopaminergic neurons by reward after the CR; and consequently, (iii) improved cortical representations of the CR. CR representations may therefore be maintained at a level instead of being degraded.

The effects of the cholinergic and the dopaminergic systems on cortical plasticity differ in many respects. Although dopaminergic activity may enhance or reduce cortical representation depending on the stimulus contingency, the extent and direction of the cholinergic effects are largely determined by the spectral and temporal characteristics of the paired sensory stimulus (26). These differences suggest that although the ventral tegmental dopaminergic activity may be essential in contingency-based associative learning, the cholinergic system could be more engaged in stimulus feature-directed perceptual learning.

We thank Dr. D. Blake for comments on the manuscript. This research was supported by National Institutes of Health Grants NS-10414 and NS-34835, the Coleman Fund, the Mental Insight Foundation, and Hearing Research, Incorporated.

1. Recanzone, G. H., Schreiner, C. E. & Merzenich, M. M. (1993) *J. Neurosci.* **13**, 87–103.
2. Bakin, J. S., South, D. A. & Weinberger, N. M. (1996) *Behav. Neurosci.* **110**, 905–913.
3. Diamond, D. M. & Weinberger, N. M. (1986) *Brain Res.* **372**, 357–360.
4. Edeline, J. M. & Weinberger, N. M. (1993) *Behav. Neurosci.* **107**, 82–103.
5. Diamond, D. M. & Weinberger, N. M. (1989) *Behav. Neurosci.* **103**, 471–494.
6. Gao, E. & Suga, N. (2000) *Proc. Natl. Acad. Sci. USA* **97**, 8081–8086.
7. Ji, W., Gao, E. & Suga, N. (2001) *J. Neurophysiol.* **86**, 211–225.
8. Ohl, F. W. & Scheich, H. (1996) *Eur. J. Neurosci.* **8**, 1001–1017.
9. Disterhoft, J. F. & Olds, J. (1972) *J. Neurophysiol.* **35**, 665–679.
10. Merzenich, M. M., Recanzone, G. H., Jenkins, W. M. & Grajski, K. A. (1990) *Cold Spring Harbor Symp. Quant. Biol.* **55**, 873–887.
11. Weinberger, N. M. (1995) *Annu. Rev. Neurosci.* **18**, 129–158.
12. Irvine, D. R. & Rajan, R. (1996) *Clin. Exp. Pharmacol. Physiol.* **23**, 939–947.
13. Bear, M. F. & Singer, W. (1986) *Nature* **320**, 172–176.
14. Juliano, S. L., Ma, W. & Eslin, D. (1991) *Proc. Natl. Acad. Sci. USA* **88**, 780–784.
15. Kilgard, M. P. & Merzenich, M. M. (1998) *Science* **279**, 1714–1718.
16. Sachdev, R. N., Lu, S. M., Wiley, R. G. & Ebner, F. F. (1998) *J. Neurophysiol.* **79**, 3216–3228.
17. Webster, H. H., Hanisch, U. K., Dykes, R. W. & Biesold, D. (1991) *Somatosens. Motil. Res.* **8**, 327–346.
18. Zhu, X. O. & Waite, P. M. (1998) *Cereb. Cortex* **8**, 63–72.
19. Hollerman, J. R. & Schultz, W. (1998) *Nat. Neurosci.* **1**, 304–309.
20. Pirch, J., Rigdon, G., Rucker, H. & Turco, K. (1991) *Adv. Exp. Med. Biol.* **295**, 219–231.
21. Stark, H. & Scheich, H. (1997) *J. Neurochem.* **68**, 691–697.
22. Wilson, F. A. & Rolls, E. T. (1990) *J. Neurosci.* **10**, 1254–1267.
23. Bakin, J. S. & Weinberger, N. M. (1996) *Proc. Natl. Acad. Sci. USA* **93**, 11219–11224.
24. Bao, S., Chan, V. T. & Merzenich, M. M. (2001) *Nature* **412**, 79–83.
25. Kilgard, M. P. & Merzenich, M. M. (1998) *Nat. Neurosci.* **1**, 727–731.
26. Kilgard, M. P., Pandya, P. K., Vazquez, J., Gehl, A., Schreiner, C. E. & Merzenich, M. M. (2001) *J. Neurophysiol.* **86**, 326–338.
27. Kiskey, M. A. & Gerstein, G. L. (2001) *Eur. J. Neurosci.* **13**, 1993–2003.
28. Mercado, E., Bao, S., Orduna, I., Gluck, M. A. & Merzenich, M. M. (2001) *Neuroreport* **12**, 2283–2287.
29. Shulz, D. E., Sosnik, R., Ego, V., Haidarliu, S. & Ahissar, E. (2000) *Nature* **403**, 549–553.
30. Otmakhova, N. A. & Lisman, J. E. (1998) *J. Neurosci.* **18**, 1270–1279.
31. Hasselmo, M. E. & Barkai, E. (1995) *J. Neurosci.* **15**, 6592–6604.
32. Gurden, H., Takita, M. & Jay, T. M. (2000) *J. Neurosci.* **20**, RC106:1–5.
33. Brocher, S., Artola, A. & Singer, W. (1992) *Brain Res.* **573**, 27–36.
34. Rioult-Pedotti, M. S., Friedman, D. & Donoghue, J. P. (2000) *Science* **290**, 533–536.
35. Buonomano, D. V. & Merzenich, M. M. (1998) *Annu. Rev. Neurosci.* **21**, 149–186.
36. Otani, S., Auclair, N., Desce, J. M., Roisin, M. P. & Crepel, F. (1999) *J. Neurosci.* **19**, 9788–9802.
37. Law-Tho, D., Desce, J. M. & Crepel, F. (1995) *Neurosci. Lett.* **188**, 125–128.
38. Markram, H., Lubke, J., Frotscher, M. & Sakmann, B. (1997) *Science* **275**, 213–215.
39. Zhang, L. I., Tao, H. W., Holt, C. E., Harris, W. A. & Poo, M. (1998) *Nature* **395**, 37–44.
40. Schuett, S., Bonhoeffer, T. & Hubener, M. (2001) *Neuron* **32**, 325–337.
41. Yao, H. & Dan, Y. (2001) *Neuron* **32**, 315–323.
42. Siegel, S. & Domjan, M. (1971) *Learn. Motiv.* **2**, 1–11.
43. Condon, C. D. & Weinberger, N. M. (1991) *Behav. Neurosci.* **105**, 416–430.
44. Chowdhury, S. A. & Suga, N. (2000) *J. Neurophysiol.* **83**, 1856–1863.
45. Sakai, M., Kudoh, M. & Shibuki, K. (1999) *Neurosci. Res.* **33**, 87–97.
46. Thompson, R. F. & Spencer, W. A. (1966) *Psychol. Rev.* **73**, 16–43.
47. Schultz, W., Apicella, P. & Ljungberg, T. (1993) *J. Neurosci.* **13**, 900–913.
48. O'Brien, J. H. & Fox, S. S. (1969) *J. Neurophysiol.* **32**, 285–296.
49. Brons, J. F. & Woody, C. D. (1980) *J. Neurophysiol.* **44**, 605–615.
50. Nudo, R. J., Milliken, G. W., Jenkins, W. M. & Merzenich, M. M. (1996) *J. Neurosci.* **16**, 785–807.