Representation of reward feedback in primate auditory cortex

Michael Brosch*, Elena Selezneva and Henning Scheich

Leibniz Institut für Neurobiologie, Magdeburg, Germany

It is well established that auditory cortex is plastic on different time scales and that this plasticity is driven by the reinforcement that is used to motivate subjects to learn or to perform an auditory task. Motivated by these findings, we study in detail properties of neuronal firing in auditory cortex that is related to reward feedback. We recorded from the auditory cortex of two monkeys while they were performing an auditory categorization task. Monkeys listened to a sequence of tones and had to signal when the frequency of adjacent tones stepped in downward direction, irrespective of the tone frequency and step size. Correct identifications were rewarded with either a large or a small amount of water. The size of reward depended on the monkeys' performance in the previous trial: it was large after a correct trial and small after an incorrect trial. The rewards served to maintain task performance. During task performance we found three successive periods of neuronal firing in auditory cortex that reflected (1) the reward expectancy for each trial, (2) the reward-size received, and (3) the mismatch between the expected and delivered reward. These results, together with control experiments suggest that auditory cortex receives reward feedback that could be used to adapt auditory cortex to task requirements. Additionally, the results presented here extend previous observations of non-auditory roles of auditory cortex and shows that auditory cortex is even more cognitively influenced than lately recognized.

Keywords: prediction error, temporal difference error, learning, dopamine, extinction

INTRODUCTION

It is widely acknowledged that auditory cortex, like many other cortical regions, remains plastic during adulthood (e.g., Dahmen and King, 2007). Auditory cortex plasticity develops over different time scales following damage to lower stages in the auditory system (e.g., Robertson and Irvine, 1989; Rajan and Irvine, 2010), after repetitively pairing acoustic with neuromodulatory signals (Bakin and Weinberger, 1996; Kilgard and Merzenich, 1998; Bao et al., 2001), during auditory perceptual learning (Recanzone et al., 1993; Zhou et al., 2010), or during task performance and task switching (Fritz et al., 2003; Atiani et al., 2009). A prerequisite for many of these changes is the establishment of appropriate cognitive associations between auditory stimuli, behavior, and reinforcement (Blake et al., 2006), which is under control of various neuromodulatory systems (Thiel et al., 2002; Suga and Ma, 2003; Weinberger, 2007). While the conditions resulting in auditory cortex plasticity are well understood, little is known about reinforcement signals reaching auditory cortex or other sensory cortices. Reinforcement is not only required for learning new tasks but also to avoid extinction, i.e., to maintain appropriate sensory motor mappings, particularly in classically and instrumentally conditioned animals, or for selecting between such previously learned mappings. Reinforcement can be mediated both by appetitive (rewarding) and aversive stimuli.

A small number of studies have found neuronal activity in auditory cortex and other sensory cortices that is related to appetitive or aversive stimuli that are meant to act as reinforcers (Pleger et al., 2008; Serences, 2008). In animals classically conditioned by pairing an auditory (Kitzes et al., 1978; Quirk et al., 1997; Armony et al., 1998) or a visual stimulus (Rowland et al., 1985) with a foot shock or with brief electrical stimulation of the medial forebrain bundle, neuronal discharges, or local field potentials were tonically increased during the interval between the conditioned and unconditioned stimulus. Once such contingencies were abandoned the tonic activity disappeared, indicating the importance of appropriately pairing stimuli and reinforcers for learning as well as for selecting and maintaining sensory motor mappings. Comparable increases of neuronal activity were seen in instrumentally conditioned animals that had to execute a motor response after an auditory (Gottlieb et al., 1989; Shinba et al., 1995; Yin et al., 2008) or visual stimulus (Shuler and Bear, 2006). Unfortunately these experiments have not been able to unequivocally disambiguate whether the neuronal activity was related to the reinforcers or to other events, such as sensory stimuli or motor behavior. This was ruled out, for instance, in recordings from non-primary auditory thalamus (Komura et al., 2001). In that study, neuronal firing was modified when the behavioral procedure was performed with rewards of differing relative values.

The present study addresses the question of whether neuronal activity in auditory cortex reflects the reward feedback that is used to motivate a subject to perform a motor response to an auditory stimulus. To this end, we recorded neuronal discharges from the auditory cortex of monkeys instrumentally trained to perform a demanding auditory categorization task. The monkeys were required to listen to sequences of tones with variable frequencies and had to signal, by release of a touch bar, when the frequency of adjacent tones stepped in a downward direction, irrespective of the tone frequency, and step size. To be able to separate influences on neuronal activity by reward/motivation from motor-related aspects and from stimulus processing, we used a reward schedule with several reward levels and reward expectations. The reward

level depended on the momentary performance of the monkey. In contrast to the reward schedule used by Bowman et al. (1996), in which monkeys were required to complete several successful trial before a reward was given, a reward was delivered after every correct response. The standard reward-size of 0.15 ml was increased to 0.22 ml when a trial with correct behavioral response was preceded by a correct trial. Note that in this reinforcement schedule, the reward level was under the subject's behavioral control (rather than under external control), such that subjects could increase the reward rate by working more consistently on the auditory categorization task over the course of consecutive trials.

MATERIALS AND METHODS

SUBJECTS

All studies were approved by the authority for animal care and ethics of the federal state of Saxony Anhalt (No. 43.2-42502/2-502 IfN) and conformed to the rules for animal experimentation of the European Communities Council Directive (86/609/EEC). Experiments were performed on two adult male long-tailed macaque monkeys (*Macaca fascicularis*) in a double-walled soundproof room (IAC 1202-A). Throughout the experiments, the two monkeys were housed together in a cage, in which they had free access to dry food including pellets, bread, corn flakes, and nuts. They earned a large proportion of their water ration during the positive-reinforcement training sessions and received the remainder in the form of fresh fruit during and after each session. On days without behavioral testing they received water and fruit. The body weight was controlled daily and never varied more than 10% from the average.

BEHAVIORAL PROCEDURE

The monkeys were seated in a primate chair, whose front compartment accommodated a red light-emitting diode, a touch bar, and a water spout; all of which were controlled remotely by computer. The water spout was connected through a plastic tube to a magnetic valve, located outside the sound-proof room.

The training of the monkeys was divided into four phases, with increasing task difficulty (Brosch et al., 2004). Both stimulus properties and reward contingencies were adjusted carefully, and gradually during the course of the training to keep the monkeys at reasonable reward rates and, thus, in a motivated and non-frustrated state. Individual training sessions lasted between 2 and 4 h, including pauses, during which time the subjects made 300–800 trials. In phase I, subjects were trained a same/different rule for acoustic items that differed along several physical dimensions (15 sessions in monkey F and 71 sessions in monkey B). In phase II, subjects had to generalize the same/different rule for acoustic items that differed along the frequency dimension only (53 sessions in monkey F and 55 sessions in monkey B). In phase III, the ultimate task was trained and animals were required to categorize tone steps (see below). It took 199 sessions in monkey F and 211 sessions in monkey B, until a clear categorization of tone steps could be detected. In the subsequent phase IV, we continued training monkey F for another 167 sessions and monkey B for another 185 sessions on the same task. In these sessions, we used tone sequences with two (instead of one) tone step sizes and fewer tone sequences, but still covering a wide frequency range.

At the end of phase IV and during the subsequent recording sessions the monkeys were required to categorize the direction of tone steps within tone sequences **Figure 1** (see also Brosch et al., 2005; Selezneva et al., 2006). A trial started with the illumination of the cue-light, and was the signal for the monkeys to grasp a touch bar. After holding this bar for 2.22 s, a sequence of up to 11 tones started. This sequence always commenced with three tones of identical frequency (black rectangles). The frequency was varied across trials in ½-octave steps over a range of 4.5 octaves, with the tone duration and intertone intervals set at 200 ms. These tones were followed by three tones of lower frequency (open rectangles), presented either immediately or following three to five intermittent tones of higher frequency (gray rectangles). Thus, the monkeys listened either to sequences with a down-step at the fourth position, or to sequences with an up-step at the same position and a down-step at some later position. The size of the tone steps was either ½ or 1 octave. The monkeys' task was to release the touch bar upon a down-step within 240–1240 ms after the onset of a tone with a lower frequency, which resulted in the monkey being rewarded with water. The release was followed by a 6-s intertrial period in which the monkeys could consume the water. A 5-s time-out was added when the monkeys prematurely released the touch bar before (false alarm) or after (miss) the 1000-ms response window.

We used a performance-dependent reward schedule, in which the amount of reward the monkeys could earn in a trial depended on the correctness of their behavioral response in the preceding trial. The reward was large (0.22 ml water) if the monkey had responded correctly in the previous trial, and the reward was small (0.15 ml water) if the previous response was incorrect. The large reward arrived at the spout 280 ms after bar release, the small at 340 ms. In some sessions we slightly modified the standard reward schedule by selectively changing large reward trials. (1) We randomly switched between trials in which the large reward was given early (530 ms) or late (890 ms) after bar release. (2) An extra-large reward (0.29 ml) instead of the standard large reward was administered in 25% of the trials in a session.

ANIMAL PREPARATION

After completion of the behavioral training paradigm, a head holder and a recording chamber were surgically implanted into the monkeys' skull (Brosch and Scheich, 2008). These implants were required for atraumatic head restraint and for accessing the brain with electrodes. All surgical procedures were performed under deep general anesthesia followed by a full course of antibiotic (Amoxicillin, Duphamox, Fort Dodge) and analgesic (Novalgin, Aventis) treatment.

ACOUSTIC STIMULI

A computer, interfaced with an array processor (Tucker-Davis Technologies, Gainesville) was used to generate acoustic stimuli at a sampling rate of 100 kHz. The signal was D/A converted, amplified (Pioneer, A202) and fed to a free-field loudspeaker (Manger, Mellrichstadt), which was placed 1.2 m and 40° from the midline into the right side of the animal. The sound pressure level (SPL) was measured with a free-field 1/2′ microphone (40AC, G.R.A.S., Vedbak), located close to the monkey's head, and a spectrum analyzer (SA 77, Rion).

ELECTROPHYSIOLOGY

Electrophysiological recordings were performed with a seven-electrode system (Thomas Recording). Electrode impedance ranged between 2 and 4 MΩ (measured at 1 kHz). The system was oriented at an angle of ~45° in the dorsoventral plane such that electrodes penetrated the dura approximately at a right angle and either directly reached auditory cortex or first traversed parietal cortex. We only included (1) sites at which neurons responded to tones of different frequencies or to noise bursts and (2) sites that were more ventral and less than 1 mm in the supratemporal plane from a site with an auditory response. Thus, only recordings from the auditory cortex entered our analysis. Areal membership was determined by the spatial distribution of best frequency that was characteristic for primary auditory cortex and posterior belt fields (Kaas and Hackett, 2000). Recordings were made from a region extending 7 mm in the mediolateral direction in monkey B and 6 mm in monkey F, and from a region extending 7 mm in the caudomedial direction in monkey B, and 8 mm in monkey F, including primary auditory cortex in both monkeys.

Following preamplification, the signals from each electrode were amplified and filtered (0.5–5 kHz) to yield spikes. All data were recorded onto 32-channel A/D data acquisition systems (BrainWave; DataWave Technologies or Alpha-Map; Alpha–Omega). By means of the built-in spike detection tools of the data acquisition systems [threshold crossings (more than three times above the background signal) and duration of these crossings (between 50 and 295 μs)] we discriminated the action potentials of a few neurons in the vicinity of each electrode tip (termed multiunit) and stored the time stamp and the waveform of each action potential using a sampling rate of 20.833 or 50 kHz.

The action potentials from a single unit were extracted off-line from individual multiunit records using a template-matching algorithm. The template was created by calculating the average waveform from a selection of large, visually similar spike shapes. Subsequently, the waveforms of all events in a multiunit record were cross correlated with the template; thus, waveforms were considered to be generated by the same neuron when the normalized cross correlation maximum was >0.9. This separation was followed by verifying that there were no first-order interspike intervals <1.5 ms, e.g., smaller than the refractory period of single units in the cortex.

For each reward condition, we computed a peri-event time histogram (PETH) from the firing in each multiunit or single unit record using a bin-size of 50 ms (500 ms when the two types of behavioral errors were compared to account for the small number of trials), with counting triggered when the monkey released the touch bar (reward-size coding and coding of reward mismatch) or grasped it (reward-expectancy coding). In error trials with misses, the trigger was the offset of the last tone in the sequence. Reward-related effects were also detectable with other bin-sizes. The standard bin-size of 50 ms was chosen because it provided both an appropriate temporal resolution of reward effects and a reasonable power of statistical tests. We used a bootstrap procedure to determine the bins in which the PETHs of two conditions were significantly different. For each bin, we obtained the distribution of the number spikes from all trials. After pooling the observations of the two conditions, the difference in sample means was calculated and recorded for every possible way of dividing these pooled values into two conditions (i.e., for every permutation of the two conditions). The one-sided p-value of the test is calculated as the proportion of sampled permutations where the difference in means was greater than or equal to the observed difference of the two conditions.

For reward-size coding we compared trials with large, small, or no reward or trials with large and extra-large reward, or trials with different delivery times for the large reward. For reward mismatch coding, we compared correct trials in which the monkeys expected and received either a small or a large reward (zero) with false alarm trials in which the monkeys received no reward despite expecting either a small or a large reward (small or large). For expectancy coding, we compared trials that were preceded by a rewarded trial (large expectancy) with trials that were preceded by an unrewarded trial (small expectancy).

RESULTS

Out of a total of 626 multiunits recorded from two macaque monkeys during the performance of an auditory categorization task with a performance-dependent reward schedule, we observed that neuronal firing in auditory cortex reflected: (i) the reward expectancy for the upcoming trial, (ii) the size of the reward obtained in a trial, and (iii) the mismatch between the expected and the received reward in a trial (reward mismatch). No systematic differences were observed between units in primary and posterior auditory cortices. Firing related to reward-size was also seen in 74 single units.

It is likely that the monkeys were aware of the reward schedule because they performed better (77.9 vs. 73.1% in monkey F; 75.9 vs. 71.9% in monkey B; $p < 0.001$, chi-square test) and licked earlier [360 vs. 486 ms in monkey F (t-test, $p < 0.0001$); 37 vs. 44 ms in monkey B ($p < 0.05$)] in trials with large reward expectations, than they did in trials with a small expectancy. This difference
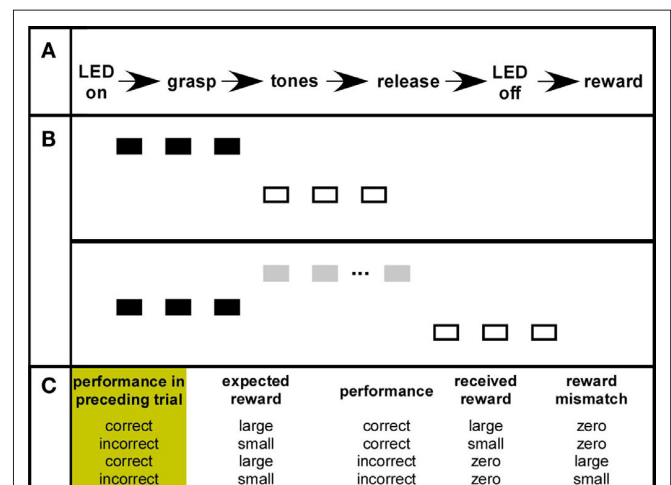

FIGURE 1 | (A) The behavioral paradigm. **(B)** Tone sequences with a downward frequency step and tone sequences with both an upward and a downward frequency step. The monkeys' task was to identify downward steps **(C)** The standard performance-dependent reward-rule. See Section "Materials and Methods" for details.

suggests that the monkeys made predictions from the outcomes of preceding trials, and did not make (probabilistic) estimates of average yield of reward.
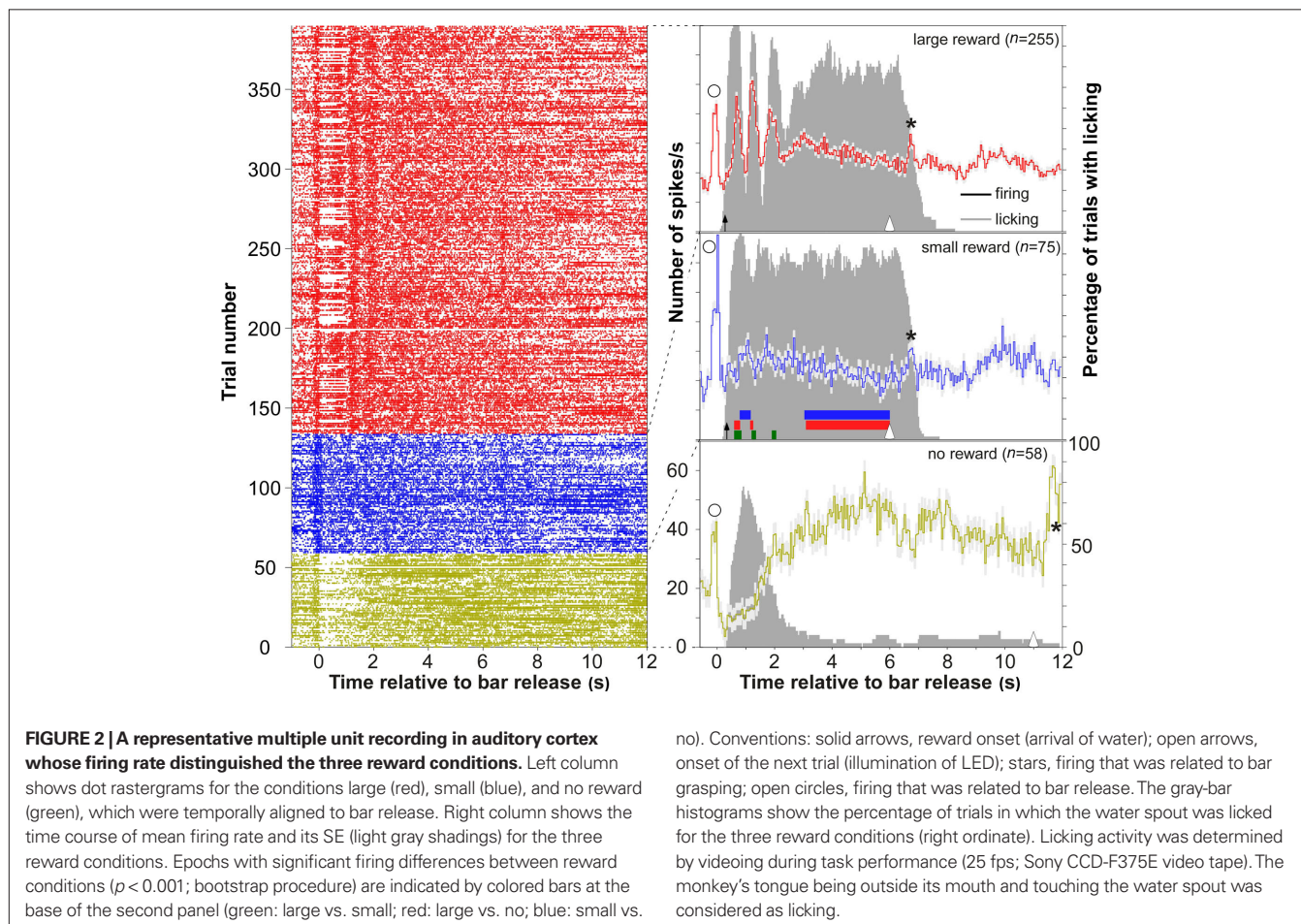
REWARD-SIZE CODING

After bar release, delivery of the reward ~300 ms later elicited neuronal firing that reflected the size of received reward. Of the 626 multiunits recorded in primary and posterior auditory cortex, 324 (51.8%) showed reward responses for a few seconds after reward delivery that discriminated reward-size by the strength of firing. A sample multiunit is shown in **Figure 2**, and the grand average of all 626 multiunits still reflecting these firing differences is shown in **Figure 3A**. When the monkey received the large reward, the firing rate increased briefly during three to four epochs. After the small reward, the periodic peaks were smaller. When the monkey received no reward for an incorrect bar release, the firing rate was slightly suppressed and significantly lower than in either of the two rewarded conditions during the first second after bar release. Firing increased slowly for ~4 s, exceeding that in the two rewarded conditions, and eventually decreased until the beginning of the next trial, 11 s after bar release in error trials. To summarize, for the first few seconds after bar release increases in firing level were related to the size of the reward, whereas later firing increased only when no reward was received.

The 324 multiunits fired significantly more spikes in at least one 50 ms bin during the intertrial period from 300 to 3000 ms after bar release ($p < 0.001$; bootstrap), when comparing the large- and small-reward conditions, the large and no-reward conditions, or the small and no-reward conditions. These differences are clearly present, even in the grand average firing of all 626 multiunits (**Figure 3A**).

In different multiunits, the increase in firing in the rewarded conditions compared with the no-reward condition was present at different times, resulting in varying percentages of active multiunits during the intertrial period, which we term "recruitment." As shown in **Figure 4A** (red curve) the percentage of recruited multiunits that coded reward-size rapidly increased to a maximum of 25.7% at 700 ms after bar release, then slowly decreasing to near zero at ~4 s. **Figure 5** shows detailed comparisons between different reward-size conditions.

When no reward was delivered, 208 multiunits (33.2%), like the multiunit in **Figure 2**, increased firing during later periods of the intertrial interval, after the initial weak or suppressed firing. These late responses almost exclusively distinguished the no-reward condition from the large- or small-reward conditions, but seldomly differentiated small from large rewards. This suggests that the late responses primarily distinguishes rewarded (correct) from unrewarded (incorrect) trials and represents a different
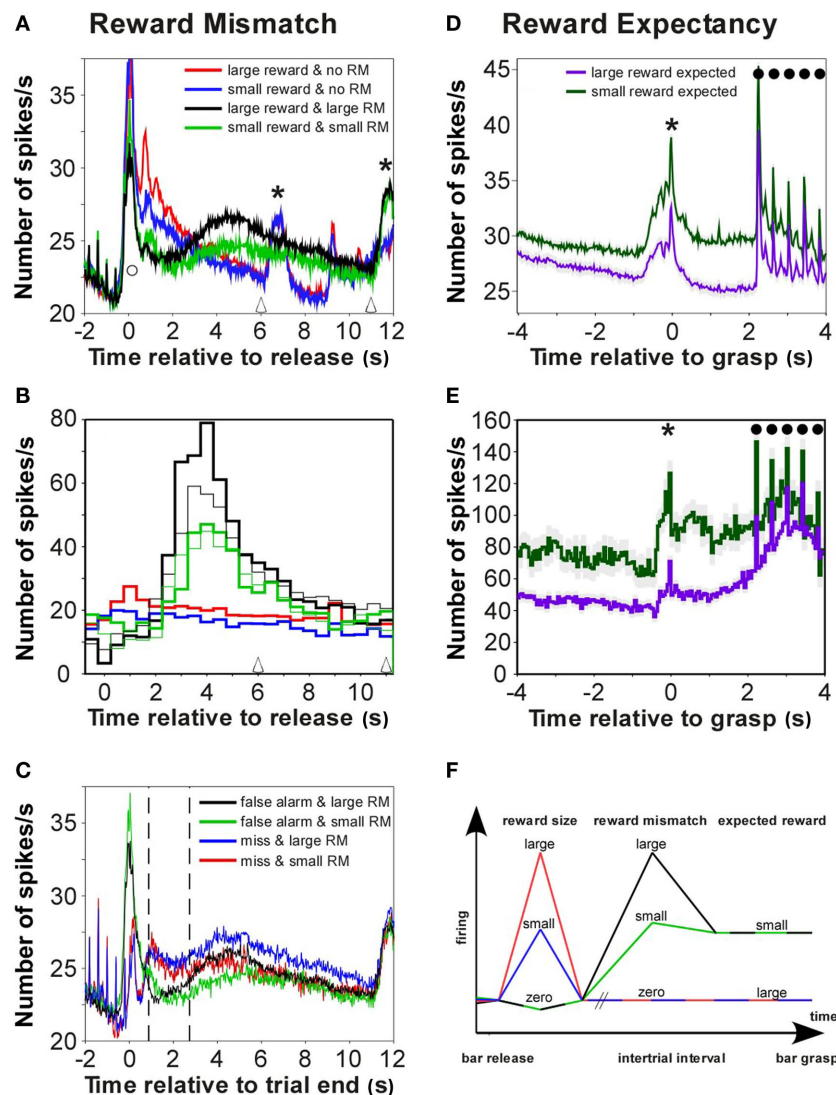

FIGURE 2 | A representative multiple unit recording in auditory cortex whose firing rate distinguished the three reward conditions. Left column shows dot rastergrams for the conditions large (red), small (blue), and no reward (green), which were temporally aligned to bar release. Right column shows the time course of mean firing rate and its SE (light gray shadings) for the three reward conditions. Epochs with significant firing differences between reward conditions ($p < 0.001$; bootstrap procedure) are indicated by colored bars at the base of the second panel (green: large vs. small; red: large vs. no; blue: small vs.

no). Conventions: solid arrows, reward onset (arrival of water); open arrows, onset of the next trial (illumination of LED); stars, firing that was related to bar grasping; open circles, firing that was related to bar release. The gray-bar histograms show the percentage of trials in which the water spout was licked for the three reward conditions (right ordinate). Licking activity was determined by videoing during task performance (25 fps; Sony CCD-F375E video tape). The monkey's tongue being outside its mouth and touching the water spout was considered as licking.
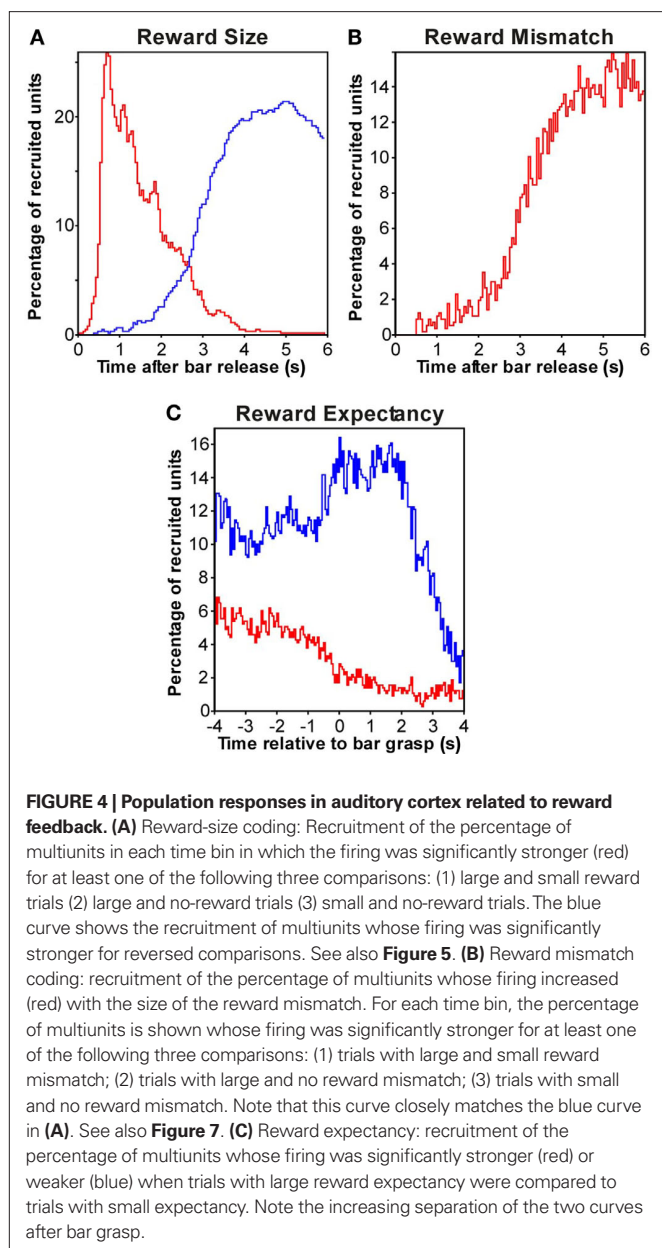
FIGURE 3 | Firing in auditory cortex related to reward expectancy and to the mismatch between expected and received reward. (A) Grand averages of the firing of 626 multiunits in auditory cortex relative to bar release for different sizes of rewards and reward mismatches (RM) between expected and received reward. The colored curves represent trials with various sizes of received rewards and subsequent mismatches: red, a large reward with no mismatch; blue, a small reward with no mismatch; black, no reward with a large mismatch; and green, no reward with a small mismatch. Note the strong firing concomitant with bar release in all cases (open circles) and the subsequent differential coding of reward-size and of the mismatch with a peak around 4 s. The next trial (open arrowheads) started earlier after correct trials than after incorrect trials. **(B)** Firing of a sample multiunit for different sizes of the reward mismatch, i.e., for different relationships between the reward expected and actually received in a trial. Conventions as in **(A)**. Thick and thin curves show error trials with false alarms or misses, respectively. **(C)** Firing in auditory cortex discriminated reward mismatches earlier in trials with misses than in trials with false alarm. In trials

with misses, turning off the cue-light and the tones indicated trial end and that no reward will become available (blue and red curves for large and small RM, respectively). In false alarm trials (like in correct trials) the cue-light and the tones were turned off immediately after bar release (black and green curves for large and small mismatches, respectively); thus there was no cue regarding whether a reward will become available. **(D)** Grand average of the reward-expectancy firing of 626 multiunits when a small reward (green) was expected, or when a large reward (blue) was expected. Filled circles indicate the responses to the tones. **(E)** A sample multiunit whose firing discriminated the size of expected reward. **(F)** Scheme of neuronal firing states in auditory cortex related to reward feedback. Early after bar release, responses distinguished large (red) from small (blue) rewards and from no rewards (black/green). Late after bar release, responses distinguished large reward mismatches (black) from small reward mismatches (green) and no reward mismatches (red/blue). Reward-expectancy firing distinguished trials in which monkeys expected a large (red/blue) reward from those in which monkeys expected a small (black/green) reward.

aspect of reward-related coding; namely the mismatch between the expected and received reward and thus the correctness of the mapping between the auditory stimuli and behavioral response (see below).

As shown in **Figure 4A** (blue curve) the percentage of recruited multiunits with late responses slowly increased after bar release, reaching a maximum of 21.4% ~5 s after bar release and then slowly decreasing. Like the multiunit in **Figure 2**, 47.8% of the multiunits

FIGURE 4 | Population responses in auditory cortex related to reward feedback. (A) Reward-size coding: Recruitment of the percentage of multiunits in each time bin in which the firing was significantly stronger (red) for at least one of the following three comparisons: (1) large and small reward trials (2) large and no-reward trials (3) small and no-reward trials. The blue curve shows the recruitment of multiunits whose firing was significantly stronger for reversed comparisons. See also **Figure 5**. **(B)** Reward mismatch coding: recruitment of the percentage of multiunits whose firing increased (red) with the size of the reward mismatch. For each time bin, the percentage of multiunits is shown whose firing was significantly stronger for at least one of the following three comparisons: (1) trials with large and small reward mismatch; (2) trials with large and no reward mismatch; (3) trials with small and no reward mismatch. Note that this curve closely matches the blue curve in **(A)**. See also **Figure 7**. **(C)** Reward expectancy: recruitment of the percentage of multiunits whose firing was significantly stronger (red) or weaker (blue) when trials with large reward expectancy were compared to trials with small expectancy. Note the increasing separation of the two curves after bar grasp.

with reward-size responses that emerge early after bar release also exhibited late responses, suggesting that many neurons encode different aspects of the reward at different times.

We can thus rule out that reward-size responses were solely due to sounds or to motor acts associated with the monkeys licking the water reward. Similar initial licking activities during the time of significant firing differences always occurred, independent of whether there was water on the spout, and therefore did not explain the firing decrease in the no-reward condition (**Figure 2**, gray histograms). Only the subsequent periodic structure of the licking in the rewarded conditions was reflected to some extent by the firing periodicity of the neurons. The missing correlation between initial licking and initial firing was confirmed in a control experiment on 70 multiunits by comparing reward responses for two reward delays (**Figures 6A,B**). Licking commenced during the time of bar

release, yet before arrival of the water; the subsequent firing pattern showed precisely the delays in water delivery. The encoding of the reward-size was further indicated in another control experiment on 12 multiunits that responded more strongly to an occasional extra-large reward (0.29 ml) than to the standard large reward of 0.22 ml (**Figure 6C**).

These experiments together suggest that both the start and the rate of the early reward-related firings are determined by the amount of water delivered even though some of the later firing may appear synchronized with licking; however, the mechanisms by which the reward-size was sensed remains unclear. It is possible that the reward could either be immediately seen by the monkey, or felt by its tongue on the spout. The reward-size coding cannot be confounded by reward expectancy, because neither the occasional extra-large rewards nor the different reward delays were predictable. As is shown later, a separate reward-expectancy coding with opposite sign was identified in auditory cortex, but only prior to reward delivery.

CODING OF THE MISMATCH BETWEEN EXPECTED AND RECEIVED REWARDS

As shown above late reward-related firing emerged only in trials in which the monkeys did not receive a reward. Thus this firing could serve as feedback signal used to inform the auditory cortex of erroneous sensory processing or erroneous sensori-motor mappings. The following considerations indicate that such error coding is mixed with the coding of the magnitude of the mismatch between the reward received in a trial and that expected for the trial.

Firing that reflected the magnitude of the mismatch between the expected and received reward is exemplified by the sample multiunit and by the grand average firing of 626 multiunits (**Figures 3A,B**). About 2 s after bar release neurons fired significantly more spikes ($p < 0.001$; bootstrap) when the difference between the expected and received reward was large (solid black curves), than when this difference was small (green curves) or zero (red and blue curves). Significantly stronger firing was also seen when the reward mismatch was small rather than zero. **Figure 7** shows more comparisons between conditions with different reward mismatches. In total, 167 (26.7%) of the multiunits exhibited firing patterns that reflected the magnitude of the mismatch between the expected and received reward.

The percentage of recruited multiunits whose firing discriminated the magnitude of the reward mismatch slowly rose after bar release, and reached a maximum of 16% after 5 s (**Figure 4B** for false alarms and **Figure 7C** for misses). Subsequently, the percentage of recruited multiunits slowly decreased within 5 s and approached zero shortly before the beginning of the next trial (**Figure 7D**). This was revealed by comparing error trials with an extended intertrial period of 11 s instead of 6 s. Late firing that related to the absence of a reward was present after different types of errors, false alarms and misses, but increased earlier in the former than in the latter (**Figure 3C**). This might be because in trials with misses, turning off both the tone sequence and the cue-light provided a cue to the monkeys that the ongoing trial was aborted, and no reward would become available.

We could rule out that late reward-related firing reflected information that was based on directly comparing the reward received in a trial with that received in the preceding trial. With analogy
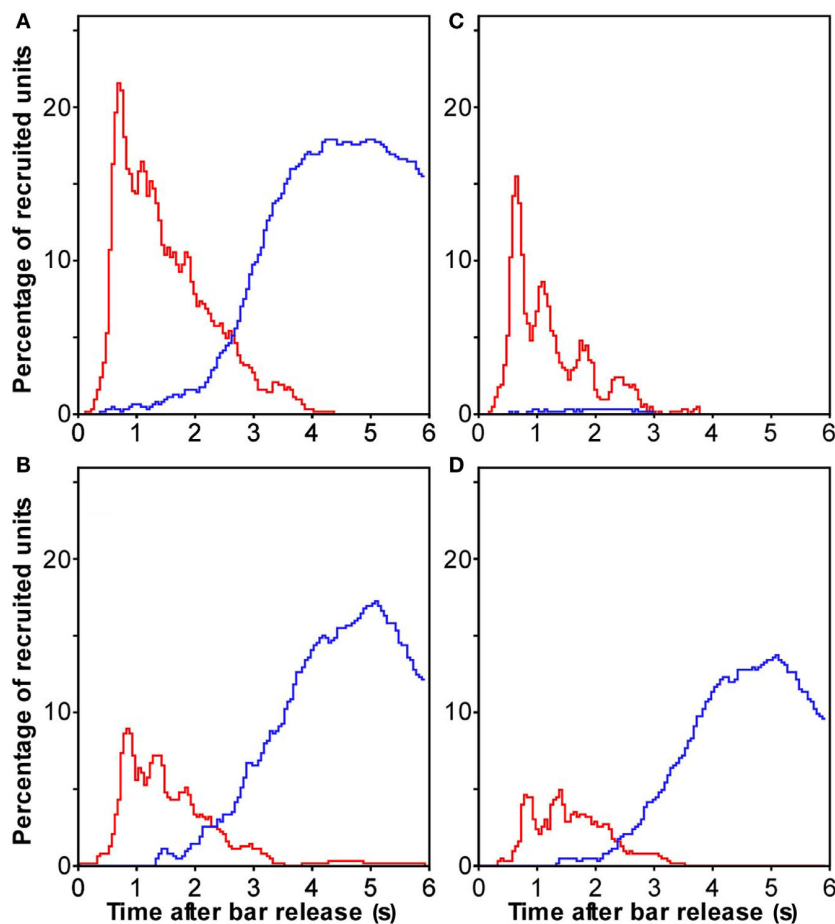
FIGURE 5 | Population responses in auditory cortex that discriminated specific reward-size conditions. (A) Recruitment of the percentage of multiunits whose firing was stronger (red) in the large reward condition than in the no-reward condition. The blue curves here and in the other panel show recruitment when the condition with the smaller reward yielded stronger

firing. Conventions as in **Figure 4**. **(B)** The same comparison for the small and no-reward conditions. **(C)** The same comparison for the large and small reward conditions. **(D)** Multiunits whose firing differed both between the large and no-reward condition and between the small and no-reward condition.
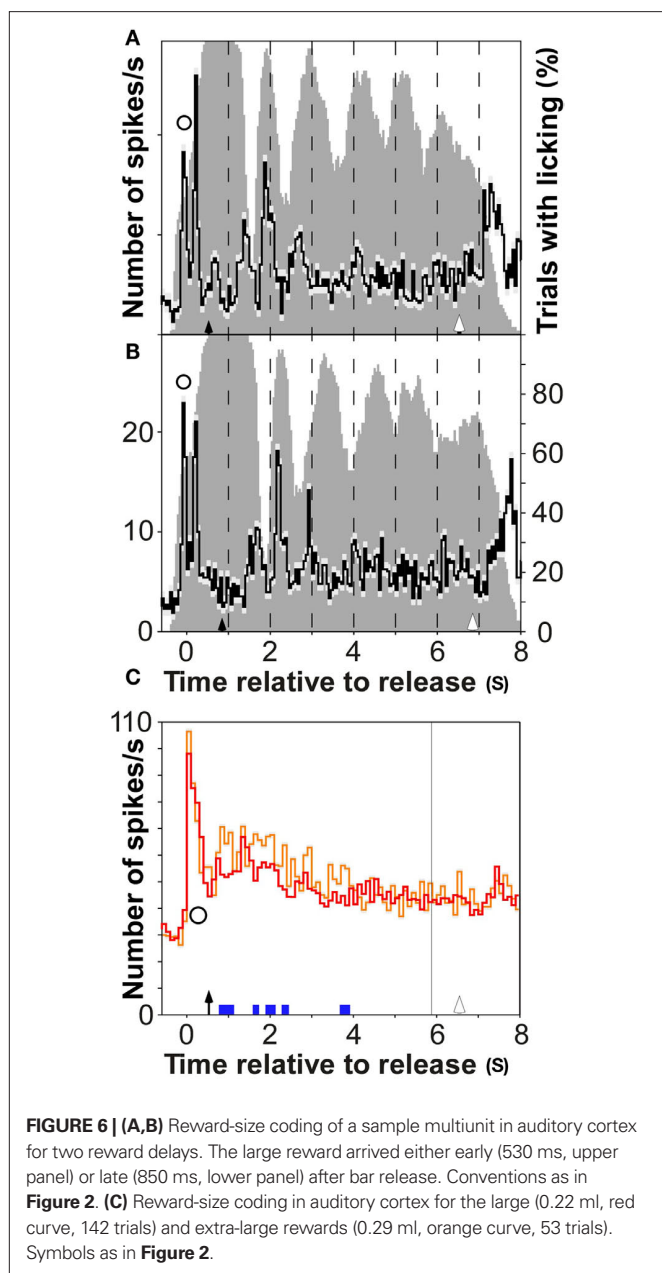
to findings in dopaminergic neurons (Schultz, 2007), we hypothesized that the reward for the preceding trial was memorized such that any change of reward led to a change in firing. Sorted in this way, late responses only partially support this scheme (**Figure 8**). As expected, late responses were not observed for two successive large rewards, but were present when a large or a small reward was followed by no reward. Contrary to the hypothesis, no late responses occurred when a small reward was followed by a large reward, or when no reward was followed by a small reward, i.e., when the reward increased in size. Also contrary to the hypothesis, late responses did occur in two successive trials with no rewards.

REWARD-EXPECTANCY CODING
Because late firing after bar release coded the magnitude of the mismatch between the expected and received reward, we searched for coding of reward expectancy in the neuronal firing relative to the beginnings of high- and low-expectation trials, using grasping of the touch bar as the reference for neuronal activity.

A total of 303 (48.4%) multiunits exhibited firing that reflected the two sizes of expected rewards, for a median duration of 750 ms from 4 s before to 4 s after bar grasp. Most (241 or 79.5%) fired more strongly when the small reward was expected, than they did when the large reward was expected (see the firing of all 626 multiunits in **Figure 3D** and the representative multiunit in **Figure 3E**). Only 20.5% exhibited the opposite relationship. The firing of the multiunit shown in **Figure 3E** was strong when the monkey scored incorrectly in the preceding trial, i.e., had received no reward and, thus could expect a small reward in the ongoing trial (green curve). The firing was significantly weaker ($p < 0.001$; bootstrap) when the monkey had scored correctly in the preceding trial, i.e., it had received either a large or small reward, thus could expect a large reward in the ongoing trial (blue curve).

The high firing during the expectation of small rewards implies that the high firing level after an incorrect unrewarded trials continues into the next trial. Conversely, low firing after a correct (rewarded) trial continued into the next trial with a large reward expectation. The percentage of multiunits with stronger

FIGURE 6 | (A,B) Reward-size coding of a sample multiunit in auditory cortex for two reward delays. The large reward arrived either early (530 ms, upper panel) or late (850 ms, lower panel) after bar release. Conventions as in **Figure 2**. **(C)** Reward-size coding in auditory cortex for the large (0.22 ml, red curve, 142 trials) and extra-large rewards (0.29 ml, orange curve, 53 trials). Symbols as in **Figure 2**.

categorical neuronal response to the decisive tone step (from the third to the fourth tone, which occurred 1.2 s after tone sequence onset; Selezneva et al., 2006) was unaffected by any preceding reward-expectancy coding and therefore presumably of purely sensory nature.

SINGLE UNIT RECORDINGS

The activity of clearly isolated single unit could be analyzed at 74 of the 626 sites at which multiunit activity was recorded from. These single units exhibited early reward-size responses (**Figures 9A,B**). However, no late responses after unrewarded behavioral responses were seen in these single units, which also displayed no systematic population relationship with the magnitude of the mismatch between the expected and delivered reward (**Figure 9B**). Additionally, single units did not show a distinction between trials with high and low reward expectancy (**Figure 9C**). We speculate that a possible explanation for different results in single units and multiunits with respect to late firing might be that preferentially those neurons in auditory cortex exhibit late and long lasting responses that have small action potentials and that are therefore less frequently isolated in standard extracellular microelectrode recordings. A similar difference between single unit and multiunit activity was also seen in our previous report for phasic and sustained firing in auditory cortex that was related to auditory and non-auditory events of the behavioral procedure (Brosch et al., 2005). While the phasic responses were observed both in single unit and multiunit activity (although with different proportions), sustained increases of firing were observed in multiunit activity only. Only two single units appeared to have such firing increases, but they were not statistically significant.

DISCUSSION

This study clearly demonstrates that the firing of neurons in auditory cortex represents different aspects of the reward feedback that is used to motivate monkeys to perform an auditory categorization task. Using a performance-dependent reward schedule with two reward levels, it was observed that shortly after bar release the firing rate varied with the magnitude of the delivered reward. A few seconds later, the firing not only distinguished rewarded from unrewarded trials, but also the magnitude of the mismatch between the expected and delivered reward. Subsequently, the firing distinguished high and low reward expectancy. These observations indicate that auditory cortex receives information about rewarding events which could be involved in adjusting the auditory cortex to current task requirements, like maintaining specific stimulus motor mappings or selecting between such different previously learned mappings.

We speculate that a key to understanding the reward-related neuronal firing in the auditory cortex in the current study is the demands of the behavioral task used in our experiments. The first element is a Pavlovian-like conditioning; the monkeys must learn that downward steps in a series of tones predict reward and later recognize them. The neuronal responses to downward tone steps become stronger than the responses to non-rewarded upward tone steps (Selezneva et al., 2006), being similar to reward-predicting responses seen in Pavlovian conditioning (Schultz, 2007). However, the reward-related task differs from Pavlovian conditioning in several essential aspects; firstly, the association
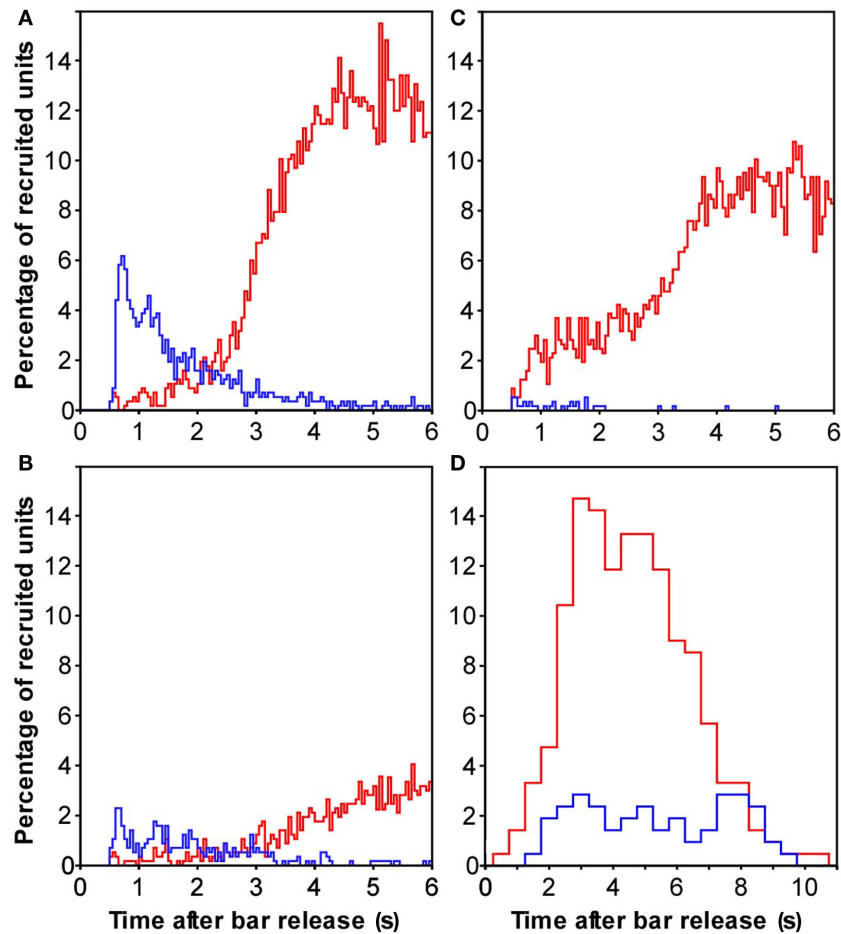
firing when small rewards were expected was at a constant level of ~11% until trial onset. After bar grasping, the percentage rose to a maximum of 16%, remaining high during the 2.2-s hold period and decreasing sometime after the onset of the tone sequence (**Figure 4C**).

In most recordings, the tone-evoked firing was superimposed on reward-expectancy related firing (**Figures 3D,E**), so we examined the end of this firing in a subgroup of multiunits that did not display additional phasic tone-evoked responses ($n = 40$; Selezneva et al., 2006). Their reward-expectancy related firing disappeared <1 s after onset of the tone sequence, rather than continuing until reward delivery >1 s later. This suggests that reward-size coding was not directly influenced by reward-expectancy coding. It also suggests that the previously described

FIGURE 7 | Population responses in auditory cortex that discriminated specific reward mismatch conditions. (A) Recruitment of multiunits whose firing was stronger (red) or weaker (blue) when trials with a large mismatch between the expected and delivered reward were compared to trials with no such mismatch. **(B)** Corresponding comparison of a small reward mismatch and no reward mismatch. **(C)** Reward mismatch coding, as shown in **Figure 4B**, except that error trials with misses instead of false alarms were used. **(D)** Recruitment of multiunits whose firing differed between reward conditions with a large and a small reward mismatch. These data were analyzed with the larger bin-size of 500 ms to account for the small number of the two types of error trials.

between the tone stimuli and reward is indirect and secondly, it depends on the choice and timely execution of an appropriate behavior; both of which are prone to mistakes. The decisive factor controlling learning is the reward feedback in response to variable behaviors that determine which of the tone steps predicts a reward. This provides a rationale for why the representation and analysis of the reward in the current task has three distinct steps: reward-size representation, reward mismatch, and reward expectancy (**Figure 3F**). The conjunction of these steps is noteworthy as it implies a type of stepwise inductive logic. By systematically monitoring how rewards change across many trials, some changes in the reward become generally predictable (obey a rule). As these changes show perseverance (i.e., they cannot be influenced), they can be ignored; whereas unpredictable changes are highlighted and clearly identify the animals' behavioral mistakes or other changes of reward supply.

The reward-related activity we observed in auditory cortex differs in several respects from neuronal activity that has previously been observed in sensory cortex and in brain structures implicated in reward processing (Schultz, 2006, 2007; Schultz and
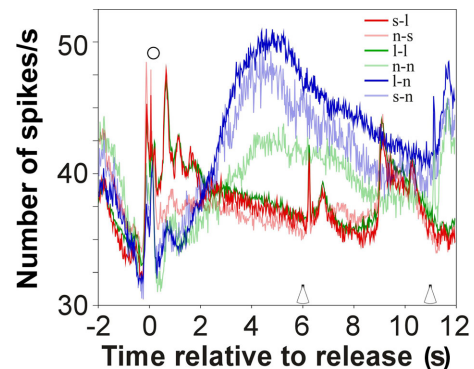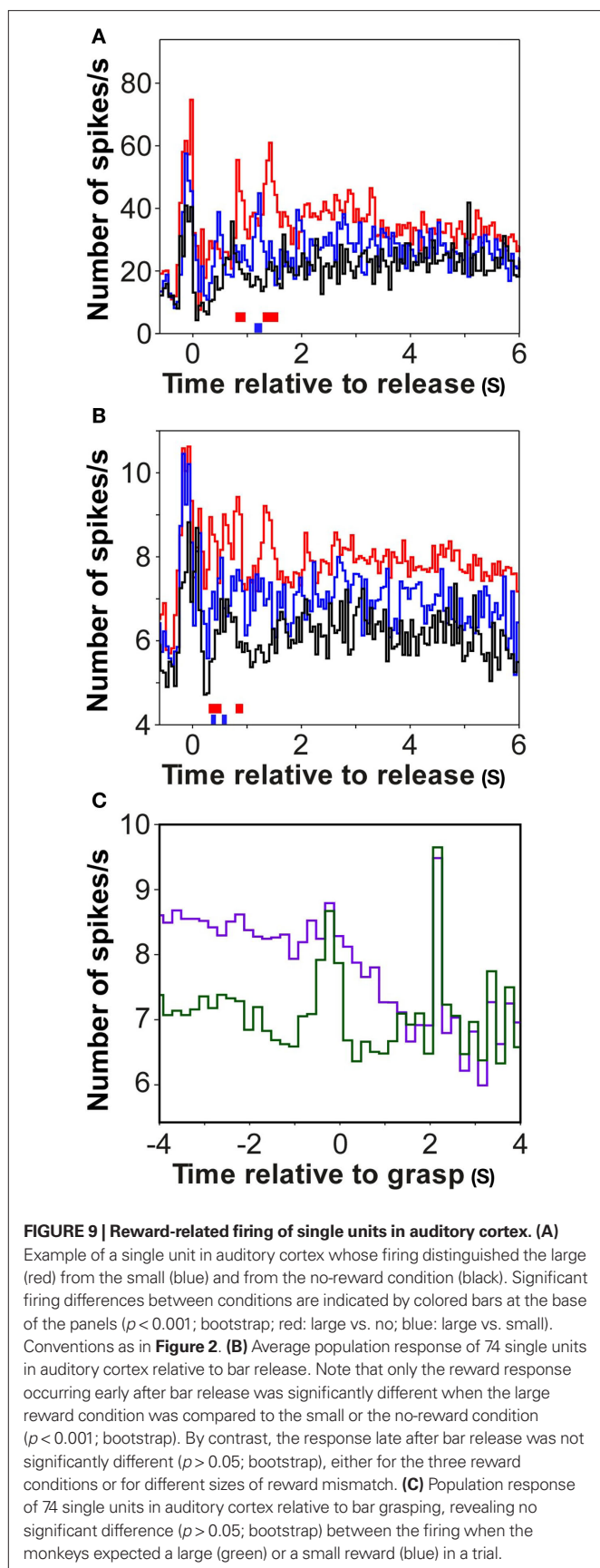

FIGURE 8 | Grand average response in auditory cortex for six relationships between the reward in a trial and that in the preceding trial. Reward increases: red (small reward followed by large reward) and pale red (no reward followed by small reward); no reward changes: green (large reward followed by large reward) and pale green (no reward followed by no reward); reward decreases: blue (large reward followed by no reward) and pale blue (small reward followed by no reward). Symbols as in **Figure 2**.

FIGURE 9 | Reward-related firing of single units in auditory cortex. (A) Example of a single unit in auditory cortex whose firing distinguished the large (red) from the small (blue) and from the no-reward condition (black). Significant firing differences between conditions are indicated by colored bars at the base of the panels ($p < 0.001$; bootstrap; red: large vs. no; blue: large vs. small). Conventions as in **Figure 2**. **(B)** Average population response of 74 single units in auditory cortex relative to bar release. Note that only the reward response occurring early after bar release was significantly different when the large reward condition was compared to the small or the no-reward condition ($p < 0.001$; bootstrap). By contrast, the response late after bar release was not significantly different ($p > 0.05$; bootstrap), either for the three reward conditions or for different sizes of reward mismatch. **(C)** Population response of 74 single units in auditory cortex relative to bar grasping, revealing no significant difference ($p > 0.05$; bootstrap) between the firing when the monkeys expected a large (green) or a small reward (blue) in a trial.

Dickinson, 2000; Holroyd and Coles, 2002; Taylor et al., 2007). Therefore it is not clear where the reinforcement related activity in auditory cortex originates from. To our knowledge, only reward expectation has been reported to be reflected in sensory cortices, but not the magnitude of the delivered reward or the mismatch between the delivered and expected reward. During classical or instrumental conditioning with positive or negative reinforcement, long lasting changes in tonic firing emerge in both auditory (Kitzes et al., 1978; Quirk et al., 1997; Armony et al., 1998; Yin et al., 2008) and visual cortices (Rowland et al., 1985; Shuler and Bear, 2006). This firing starts after a specific external stimulus and typically increases toward and ends around the time of anticipated reinforcement. In our study, we also observed tonic firing during expectation of a reward. However, this firing was triggered by the monkeys' behavior and depended on the outcome of the previous trial. Firing increased in intensity after the monkey initiated the next trial, but vanished before the presentation of a stimulus that required a behavioral response and thus well before the anticipated time of reward. In contrast to the cited studies, we could rule out that firing related to the reward-expectancy reflected aspects of the task that differed from the reward. This is because of the use of the reward-rule trials where trials with large and small rewards required the same stimulus processing and the same behavioral response.

The coding of the magnitude of the delivered reward in auditory cortex bears some similarity with coding of primary rewards by midbrain dopaminergic neurons (Bar-Gad et al., 2003; Schultz, 2004), lateral hypothalamus (Rolls et al., 1980), pedunculopontine tegmental nucleus (Okada et al., 2009), amygdala (Nishijo et al., 1988; Nakamura et al., 1992), striatum (Bowman et al., 1996; Hassani et al., 2001), and orbitofrontal cortex (Thorpe et al., 1983; Rolls et al., 1990, 1999; Tremblay and Schultz, 1999; Hikosaka and Watanabe, 2000). Neuronal responses have relatively short latencies and are short-lasting, reflecting some basic physical characteristics of the reward. The responses in auditory cortex differ from those of midbrain dopaminergic neurons in several respects; during classical conditioning, midbrain dopaminergic neurons initially respond to an offered reward only, and only after some time develop reward-predicting responses to the conditioned stimulus while no response occurs to the reward itself. Responses to rewards reappear when the reward is omitted or delayed; in which case the firing encodes errors of these reward predictions (but see Redgrave et al., 2008); firing increases when the reward increases and decreases when the reward decreases. By contrast, neurons in auditory cortex of instrumentally trained monkeys respond only slightly more strongly to the presentation of a stimulus that is associated with a reward (a tone down-step; see Selezneva et al., 2006), yet show a vigorous response to the reward itself, irrespective of whether the reward is as large as predicted or whether it is delivered at the predicted time.

The ability of midbrain dopaminergic neurons to encode prediction errors of reward seems to be more matched to the firing in auditory cortex that emerges several seconds after reward delivery or its expected delivery time, and reflects the magnitude of the mismatch between the expected and delivered rewards. This firing, however, differs from that of midbrain dopaminergic neurons in latency and duration by one order of magnitude and by its sign. Also, the firing in auditory cortex may have a bias toward

unpredictable losses but not to gains of reward. In the control experiment with extra-large rewards (**Figure 6C**) hardly any difference in late firing was observed between trials in which monkeys received the extra-large reward instead of the large reward. The same holds for the activity of other brain structures, which have been implicated with coding of (prediction) errors; including anterior cingulate and dorsolateral prefrontal and orbitofrontal cortex (Niki and Watanabe, 1979; Rosenkilde et al., 1981; Thorpe et al., 1983; Watanabe, 1989; Tremblay and Schultz, 1999; Schultz and Dickinson, 2000). Despite these differences, the properties of late reward-related firing are compatible with the requirements of a teaching signal, according to reinforcement learning theories (Sutton and Barto, 1998). However, further tests are required to understand the effects of various reward manipulations, like non-rule based reward variations, unpredicted rewards, or reward omission. Late responses in auditory cortex also differ in both latency and duration from error signals observed in the monkey and human frontal cortex (Holroyd and Coles, 2002; Taylor et al., 2007). The closest finding is neurons in anterior cingulate and dorsolateral prefrontal cortex which responded to both behavioral mistakes and reward omissions in correct trials, though these alternatives were not distinguished (Niki and Watanabe, 1979). It should also be considered that late reward-related firing in auditory cortex may have no immediate consequences for learning or performance monitoring, but is simply related to the emotional state of the animal after the expected time of reward or the motivational drive to perform the auditory task.

The current study also shows that auditory cortex may be involved in representing rules. We show that, depending on the outcome of a trial and on the reward schedule, firing in auditory cortex reflects the reward expectancy for the upcoming trial. This, together with the representation of the reward-size received in that trial is used to obtain the mismatch between the expected and delivered reward such that rule obeying variations of rewards are separated from rule violating variations in the rewards. This suggests that auditory cortex is part of a brain system that is able to make predictions about future events (Brunia, 1999; Bar, 2007) and that auditory cortex may be significantly more involved in cognitive functions then lately recognized (e.g., Scheich et al., 2007; Brosch et al., 2011).

REFERENCES

Armony, J. L., Quirk, G. J., and LeDoux, J. E. (1998). Differential effects of amygdala lesions on early and late plastic components of auditory cortex spike trains during fear conditioning. *J. Neurosci.* 18, 2592–2601.

Atiani, S., Elhilali, M., David, S. V., Fritz, J. B., and Shamma, S. A. (2009). Task difficulty and performance induce diverse adaptive patterns in gain and shape of primary auditory cortical receptive fields. *Neuron* 61, 467–480.

Bakin, J. S., and Weinberger, N. M. (1996). Induction of a physiological memory in the cerebral cortex by stimulation of the nucleus basalis. *Proc. Natl. Acad. Sci. U.S.A.* 93, 11219–11224.

Bao, S., Chan, V. T., and Merzenich, M. M. (2001). Cortical remodelling induced by activity of ventral tegmental dopamine neurons. *Nature* 412, 79–83.

Bar, M. (2007). The proactive brain: using analogies and associations to generate predictions. *Trends Cogn. Sci.* 11, 280–289.

Bar-Gad, I., Morris, G., and Bergman, H. (2003). Information processing, dimensionality reduction and reinforcement learning in the basal ganglia. *Prog. Neurobiol.* 71, 439–473.

Blake, D. T., Heiser, M. A., Caywood, M., and Merzenich, M. M. (2006). Experience-dependent adult cortical plasticity requires cognitive association between sensation and reward. *Neuron* 52, 371–381.

Bowman, E. M., Aigner, T. G., and Richmond, B. J. (1996). Neural signals in the monkey ventral striatum related to motivation for juice and cocaine rewards. *J. Neurophysiol.* 75, 1061–1073.

Brosch, M., and Scheich, H. (2008). Tone-sequence analysis in the auditory cortex of awake macaque monkeys. *Exp. Brain Res.* 184, 349–361.

Brosch, M., Selezneva, E., Bucks, C., and Scheich, H. (2004). Macaque monkeys discriminate pitch relationships. *Cognition* 91, 259–272.

Brosch, M., Selezneva, E., and Scheich, H. (2005). Non-auditory events of a behavioural procedure activate auditory cortex of highly trained monkeys. *J. Neurosci.* 25, 6797–6806.

Brosch, M., Selezneva, E., and Scheich, H. (2011). Formation of associations in auditory cortex by slow changes of tonic firing. *Hear. Res.* 271, 66–73

Brunia, C. H. (1999). Neural aspects of anticipatory behavior. *Acta Psychol. (Amst.),* 101, 213–242.

Dahmen, J. C., and King, A. J. (2007). Learning to hear: plasticity of auditory cortical processing. *Curr. Opin. Neurobiol.* 17, 456–464.

Fritz, J., Shamma, S., Elhilali, M., and Klein, D. (2003). Rapid task-related plasticity of spectrotemporal receptive fields in primary auditory cortex. *Nat. Neurosci.* 6, 1216–1223.

Gottlieb, Y., Vaadia, E., and Abeles, M. (1989). Single unit activity in the auditory cortex of a monkey performing a short term memory task. *Exp. Brain Res.* 74, 139–148.

Hassani, O. K., Cromwell, H. C., and Schultz, W. (2001). Influence of expectation of different rewards on behavior-related neuronal activity in the striatum. *J. Neurophysiol.* 85, 2477–2489.

Hikosaka, K., and Watanabe, M. (2000). Delay activity of orbital and lateral prefrontal neurons of the monkey varying with different rewards. *Cereb. Cortex* 10, 263–271.

Holroyd, C. B., and Coles, M. G. (2002). The neural basis of human error processing: reinforcement learning, dopamine, and the error-related negativity. *Psychol. Rev.* 109, 679–709.

Kaas, J. H., and Hackett, T. A. (2000). Subdivisions of auditory cortex and processing streams in primates. *Proc. Natl. Acad. Sci. U.S.A.* 97, 11793–11799.

Kilgard, M. P., and Merzenich, M. M. (1998). Cortical map reorganization enabled by nucleus basalis activity. *Science* 279, 1714–1718.

Kitzes, L. M., Farley, G. R., and Starr, A. (1978). Modulation of auditory cortex unit activity during the performance of a conditioned response. *Exp. Neurol.* 62, 678–697.

Komura, Y., Tamura, R., Uwano, T., Nishijo, H., Kaga, K., and Ono, T. (2001). Retrospective and prospective coding for predicted reward in the sensory thalamus. *Nature* 412, 546–549.

Nakamura, K., Mikami, A., and Kubota, K. (1992). Activity of single neurons in the monkey amygdala during performance of a visual discrimination task. *J. Neurophysiol.* 67, 1447–1463.

Niki, H., and Watanabe, M. (1979). Prefrontal and cingulate unit activity during timing behaviour in the monkey. *Brain Res.* 171, 213–224.

Nishijo, H., Ono, T., and Nishino, H. (1988). Single neuron responses in amygdala of alert monkey during complex sensory stimulation with affective significance. *J. Neurosci.* 8, 3570–3583.

Okada, K., Toyama, K., Inoue, Y., Isa, T., and Kobayashi, Y. (2009). Different pedunculopontine tegmental neurons signal predicted and actual task rewards. *J. Neurosci.* 29, 4858–4870.

Pleger, B., Blankenburg, F., Ruff, C. C., Driver, J., and Dolan, R. J. (2008). Reward facilitates tactile judgments and modulates hemodynamic responses in human primary somatosensory cortex. *J. Neurosci.* 28, 8161–8168.

Quirk, G. J., Armony, J. L., and LeDoux, J. E. (1997). Fear conditioning enhances different temporal components of tone-evoked spike trains in auditory cortex and lateral amygdala. *Neuron* 19, 613–624.

Rajan, R., and Irvine, D. R. (2010). Severe and extensive neonatal hearing loss in

cats results in auditory cortex plasticity that differentiates into two regions. *Eur. J. Neurosci.* 31, 1999–2013.

Recanzone, G. H., Schreiner, C. E., and Merzenich, M. M. (1993). Plasticity in the frequency representation of primary auditory cortex following discrimination training in adult owl monkeys. *J. Neurosci.* 13, 87–103.

Redgrave, P., Gurney, K., and Reynolds, J. (2008). What is reinforced by phasic dopamine signals? *Brain Res. Rev.* 58, 322–339.

Robertson, D., and Irvine, D. R. (1989). Plasticity of frequency organization in auditory cortex of guinea pigs with partial unilateral deafness. *J. Comp. Neurol.* 282, 456–471.

Rolls, E. T., Burton, M. J., and Mora, F. (1980). Neurophysiological analysis of brain-stimulation reward in the monkey. *Brain Res.* 194, 339–357.

Rolls, E. T., Critchley, H. D., Browning, A. S., Hernadi, I., and Lenard, L. (1999). Responses to the sensory properties of fat of neurons in the primate orbitofrontal cortex. *J. Neurosci.* 19, 1532–1540.

Rolls, E. T., Yaxley, S., and Sienkiewicz, Z. J. (1990). Gustatory responses of single neurons in the caudolateral orbitofrontal cortex of the macaque monkey. *J. Neurophysiol.* 64, 1055–1066.

Rosenkilde, C. E., Bauer, R. H., and Fuster, J. M. (1981). Single cell motivated bar press feeding behavior in the monkey. Brain Res. activity in ventral prefron-

tal cortex of behaving monkeys. *Brain Res.* 413, 302–313.

Rowland, V., Gluck, H., Sumergrad, S., and Dines, G. (1985). Slow and multiple unit potentials in trace and temporal conditioning controlled by electrical reward in the rat. *Electroencephalogr. Clin. Neurophysiol.* 61, 559–568.

Scheich, H., Brechmann, A., Brosch, M., Budinger, E., and Ohl, F. W. (2007). The cognitive auditory cortex, task-specificity of stimulus representations. *Hear. Res.* 229, 213–224.

Schultz, W. (2004). Neural coding of basic reward terms of animal learning theory, game theory, microeconomics and behavioural ecology. *Curr. Opin. Neurobiol.* 14, 139–147.

Schultz, W. (2006). Behavioural theories and the neurophysiology of reward. *Annu. Rev. Psychol.* 57, 87–115.

Schultz, W. (2007). Multiple dopamine functions at different time courses. *Annu. Rev. Neurosci.* 30, 259–288.

Schultz, W., and Dickinson, A. (2000). Neuronal coding of prediction errors. *Annu. Rev. Neurosci.* 23, 473–500.

Selezneva, E., Scheich, H., and Brosch, M. (2006). Dual time scales for categorical decision making in auditory cortex. *Curr. Biol.* 16, 2428–2433.

Serences, J. T. (2008). Value-based modulations in human visual cortex. *Neuron* 60, 1169–1181.

Shinba, T., Sumi, M., Iwanami, A., Ozawa, N., and Yamamoto, K. (1995). Increased neuronal firing in the rat auditory cor-

tex associated with preparatory set. *Brain Res. Bull.* 37, 199–204.

Shuler, M. G., and Bear, M. F. (2006). Reward timing in the primary visual cortex. *Science* 311, 1606–1609.

Suga, N., and Ma, X. (2003). Multiparametric corticofugal modulation and plasticity in the auditory system. *Nat. Rev. Neurosci.* 4, 783–794.

Sutton, R., and Barto, A. (1998). *Reinforcement Learning.* Cambridge: MIT Press.

Taylor, S. F., Stern, E. R., and Gehring, W. J. (2007). Neural systems for error monitoring: recent findings and theoretical perspectives. *Neuroscientist* 13, 160–172.

Thiel, C. M., Friston, K. J., and Dolan, R. J. (2002). Cholinergic modulation of experience-dependent plasticity in human auditory cortex. *Neuron* 35, 567–574.

Thorpe, S. J., Rolls, E. T., and Maddison, S. (1983). The orbitofrontal cortex: neuronal activity in the behaving monkey. *Exp. Brain Res.* 49, 93–115.

Tremblay, L., and Schultz, W. (1999). Relative reward preference in primate orbitofrontal cortex. *Nature* 398, 704–708.

Watanabe, M. (1989). The appropriateness of behavioral responses coded in post-trial activity of primate prefrontal units. *Neurosci. Lett.* 101, 113–117.

Weinberger, N. M. (2007). Auditory associative memory and representational

plasticity in the primary auditory cortex. *Hear. Res.* 229, 54–68.

Yin, P., Mishkin, M., Sutter, M., and Fritz, J. B. (2008). Early stages of melody processing: stimulus-sequence and task-dependent neuronal activity in monkey auditory cortical fields A1 and R. *J. Neurophysiol.* 100, 3009–3029.

Zhou, X., de Villers-Sidani, E., Panizzutti, R., and Merzenich, M. M. (2010). Successive-signal biasing for a learned sound sequence. *Proc. Natl. Acad. Sci. U.S.A.* 107, 14839–14844.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.