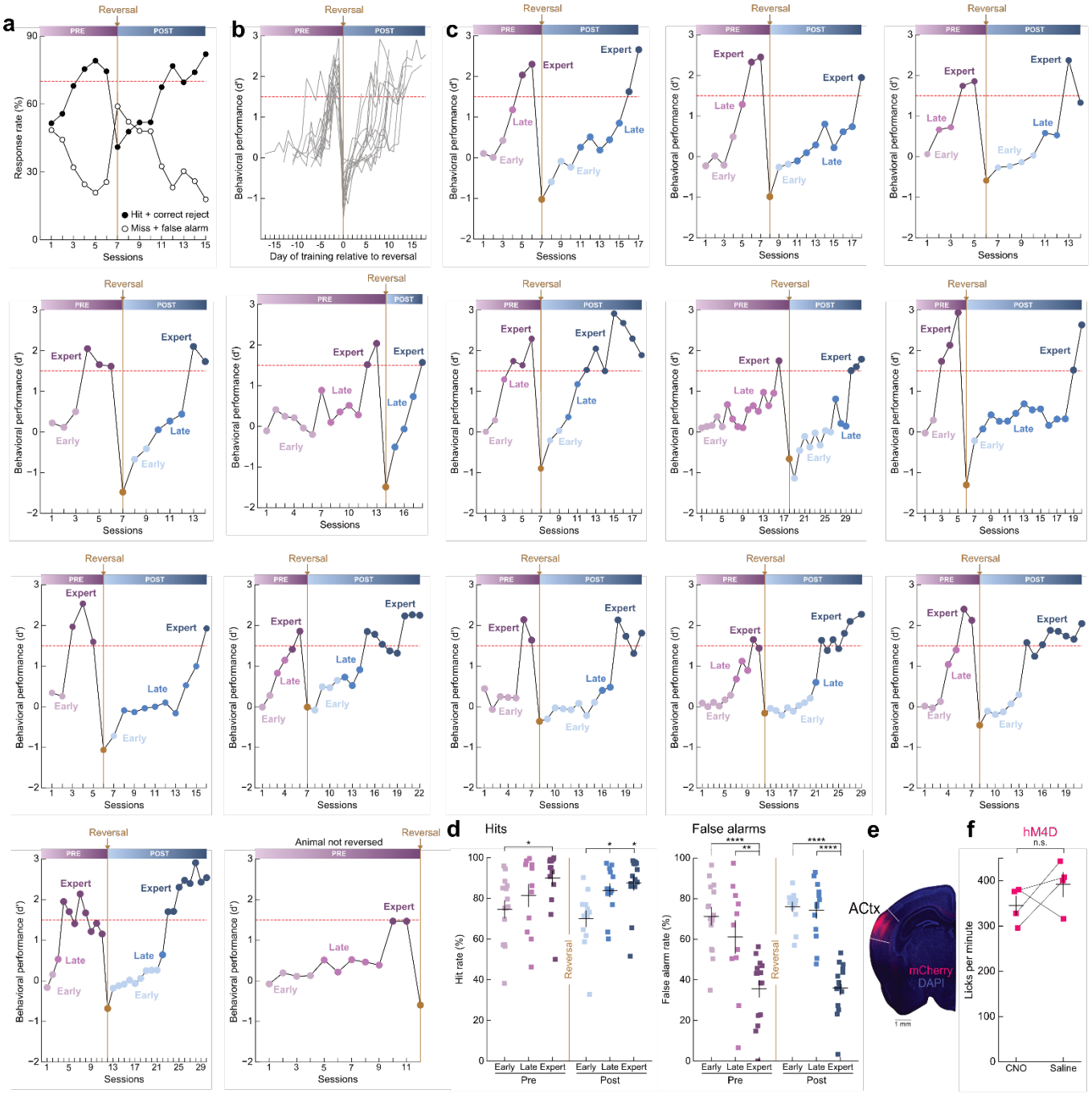
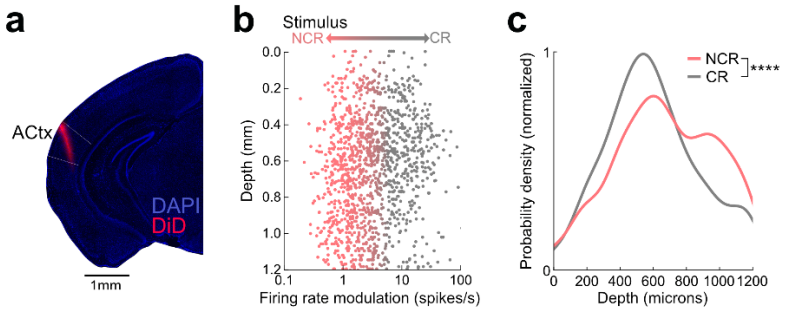
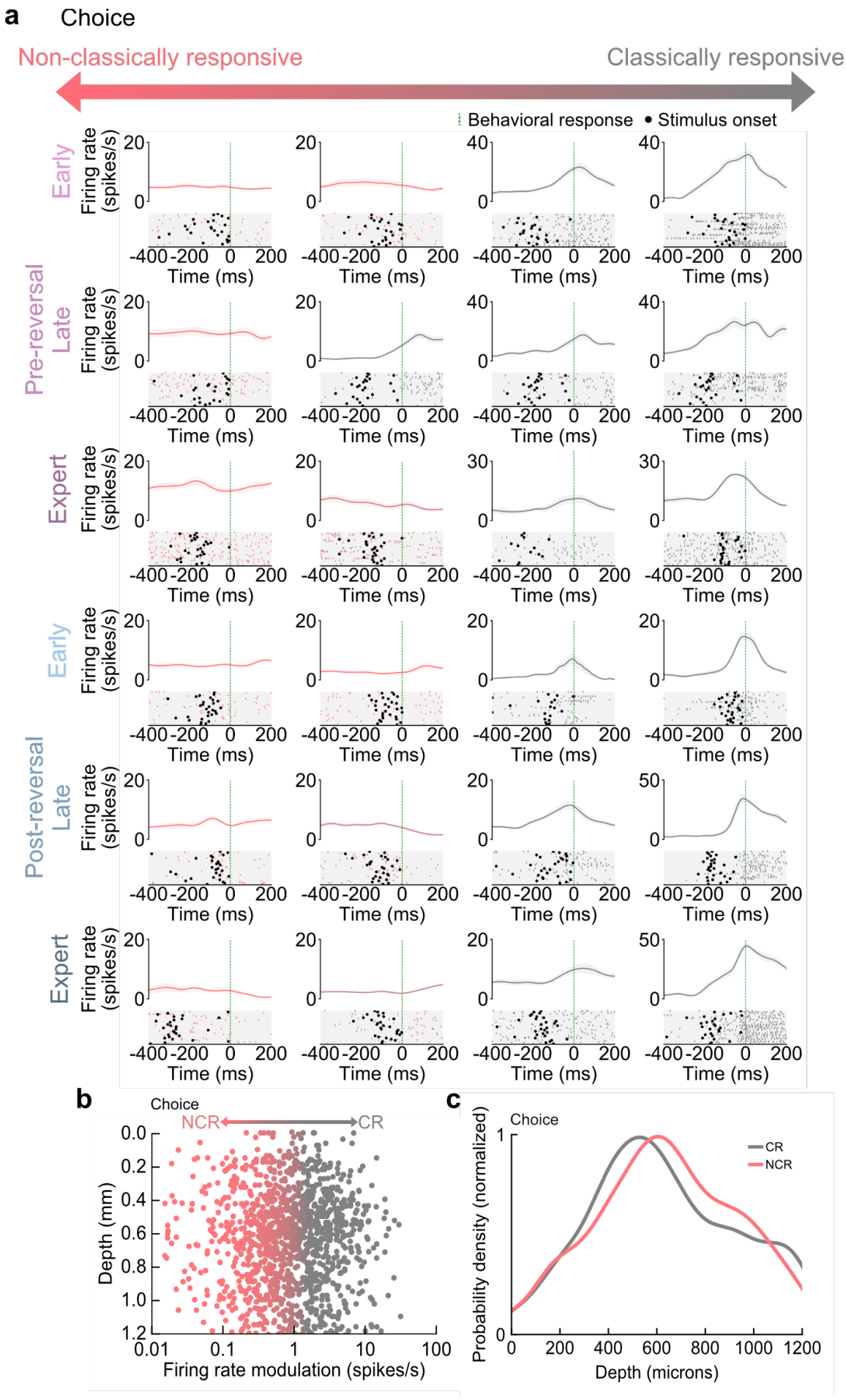


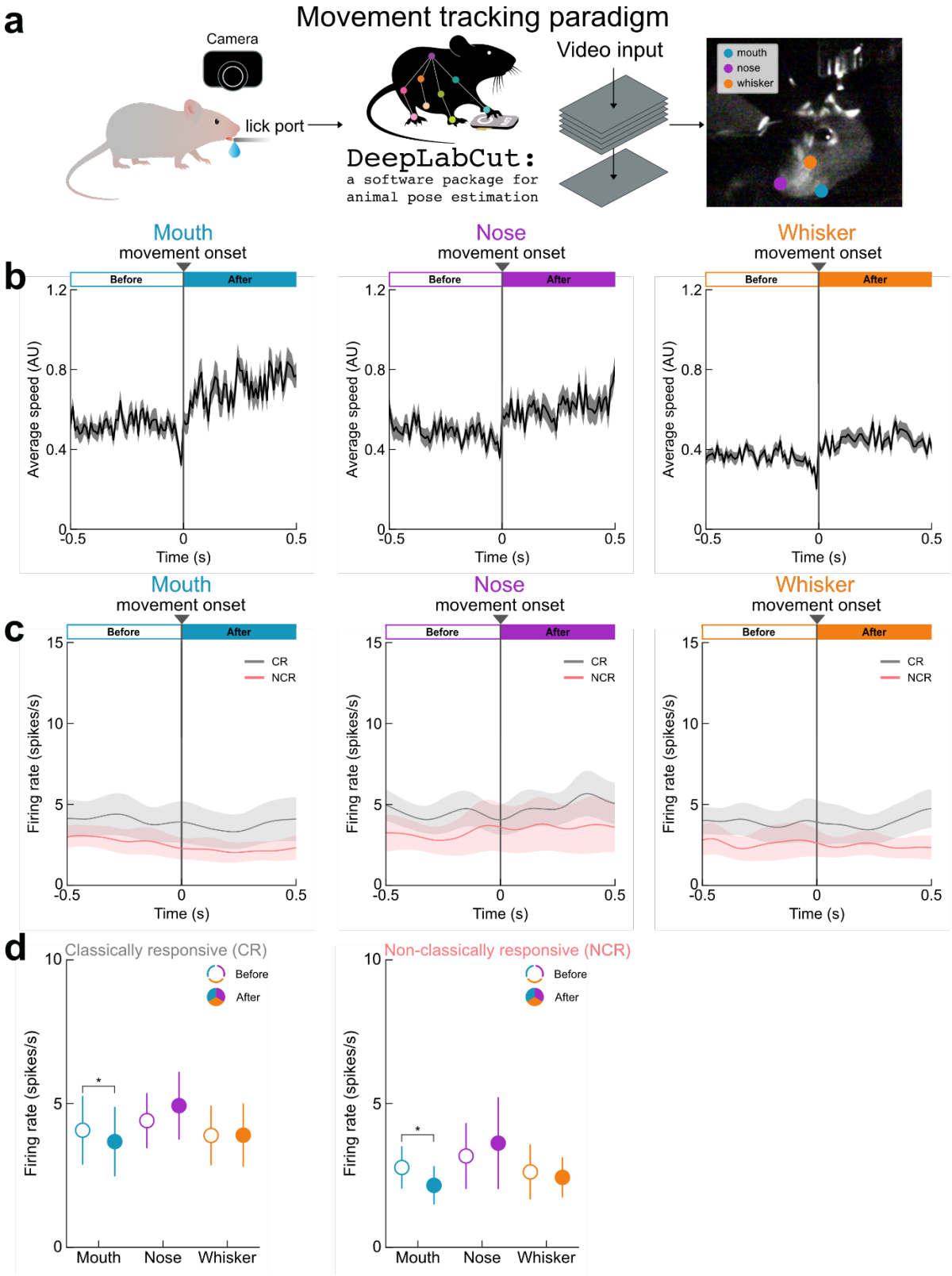
Extended Data Figure 1. Learning curves and trial outcomes from animals included in this study. a, Example trial outcome rate across reversal learning for one animal. **b**, Summary learning curves with behavioral d' values for each animal (N=15 mice). **c,** Individual learning curves from 15 animals included in this study. Each panel shows behavioral d' values across all reversal learning behavioral sessions. Early, late, and expert phases are indicated on each learning curve. **d**, Summary of trial outcome rates during each reversal learning phase. Left, hit rates. Pre_{early} vs. Pre_{expert}, $p = 0.02$; $Post_{early}$ vs. $Post_{late}$, $p = 0.02$; vs. $Post_{expert}$, $p = 0.01$. Right, false alarm rates. Pre_{expert} vs. Pre_{early}, $p = 5x10^{-6}$; vs. Pre_{late}, $p = 0.008$; $Post_{expert}$ vs. $Post_{early}$, $p = 6x10^{-9}$; vs. $Post_{late}$, $p = 4x10^{-7}$; two-sided independent t-test with Benjamini-Hochberg correction. **e**, Immunohistochemical confirmation of inhibitory DREADDs expression in primary auditory cortex. Scale bar, 1 mm. **f,** Lick rate from chemogenetically inhibited and control animals that were allowed to lick freely for water without tone presentation. Chemogenetic inactivation does not affect the ability to lick for water. CNO (N = 4 mice) vs. Saline (N = 4 mice), $p = 0.215$; two-sided independent t-test. Data are mean+s.e.m. *p<0.05, **p<0.01, ****p<0.0001.



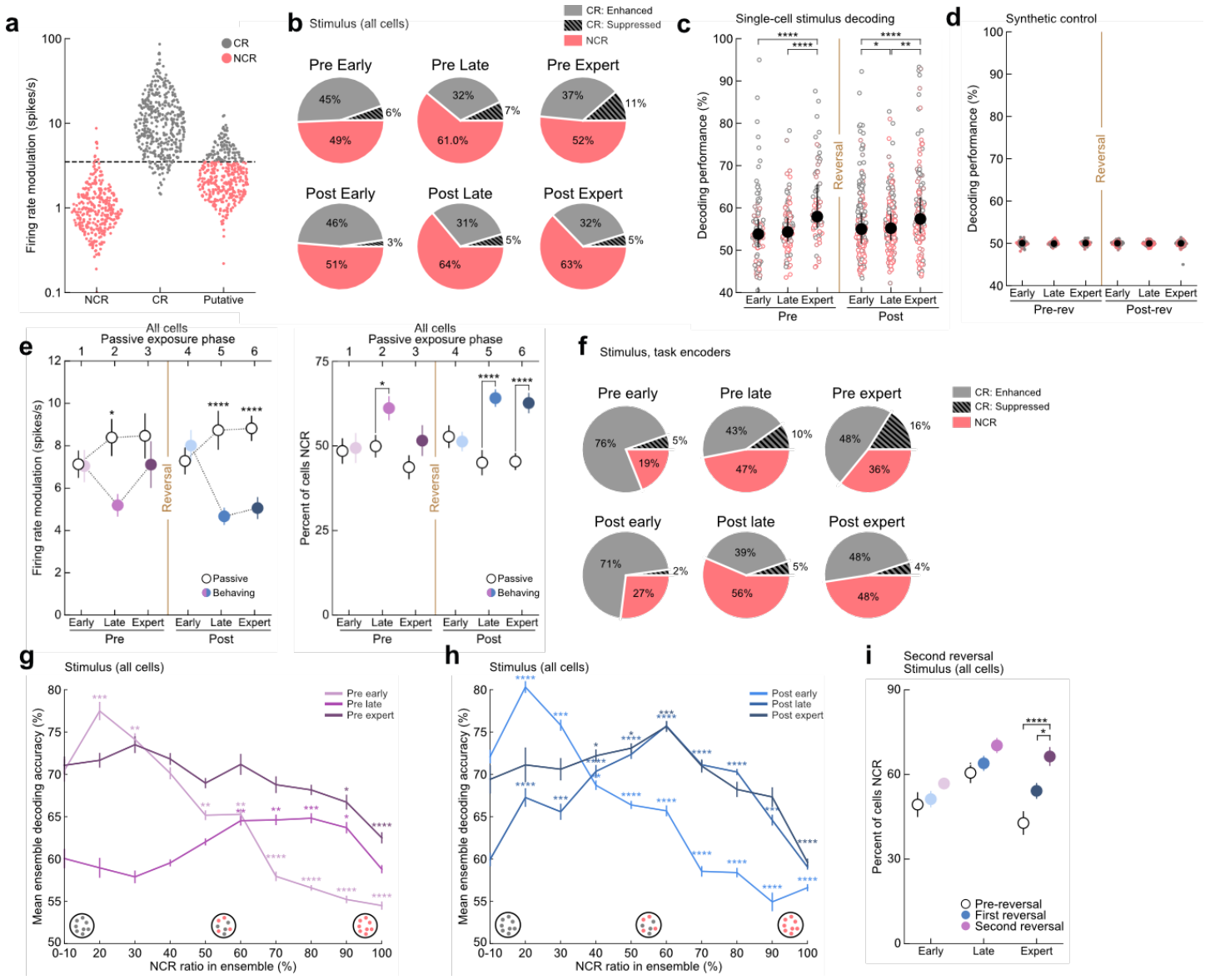
Extended Data Figure 2. Laminar distribution of heterogeneous stimulus-related cortical responses. a, Immunohistochemical confirmation of silicon probe recording location from primary auditory cortex. Probe was coated in DiD (red fluorescence) prior to recording. Scale bar, 1 mm. **b**, Laminar distribution of all recorded cells (n = 1,327 cells, N = 15 mice) as a function of the stimulus firing rate modulation; firing rate modulation vs. depth, r = -0.15, p = $3x10^{-8}$; two-sided Spearman's rank correlation test. **c**, Depth distributions of stimulus classically responsive cells (CR, gray, n = 559 cells) and non-classically responsive cells (NCR, red, n = 768 cells) shown in **b**, p = $2x10^{-6}$; two-sided Mann-Whitney U test. ****p<0.0001.


a
Choice
Non-classically responsive
Classically responsive
Behavioral response
Stimulus onset
Pre-reversal
Post-reversal
Early
Late
Expert
Firing rate (spikes/s)
Time (ms)
b
Choice
NCR
CR
Depth (mm)
Firing rate modulation (spikes/s)
c
Choice
CR
NCR
Probability density (normalized)
Depth (microns)

Extended Data Figure 3. Heterogeneous choice-related cortical responses during learning. a, Rasters and peri-stimulus time histograms (PSTHs) for 24 cortical neurons exemplifying the range of choice responses from non-classically responsive (red, NCR) to classically responsive (gray, CR) during all stages of reversal learning. Trials aligned to behavioral responses (black filled circle, stimulus onset; green dashed line, behavioral response). Lines in PSTH, mean firing rate; shading, s.e.m. For rasters, a subset of trials are shown for clarity. **b**, The range of choice-related neural responses including classically and non-classically responsive choice cells were recorded in superficial, middle, and deep cortical layers and no laminar specific topography was observed. Laminar distribution of all recorded cells (n = 1,300 cells) as a function of choice firing rate modulation; choice firing rate modulation vs. depth, $r = -0.01$, $p = 0.6$; two-sided Spearman's rank correlation test. **c**, Depth distributions of choice classically and non-classically responsive cells shown in **b**, $p = 0.4$; Mann-Whitney U test.

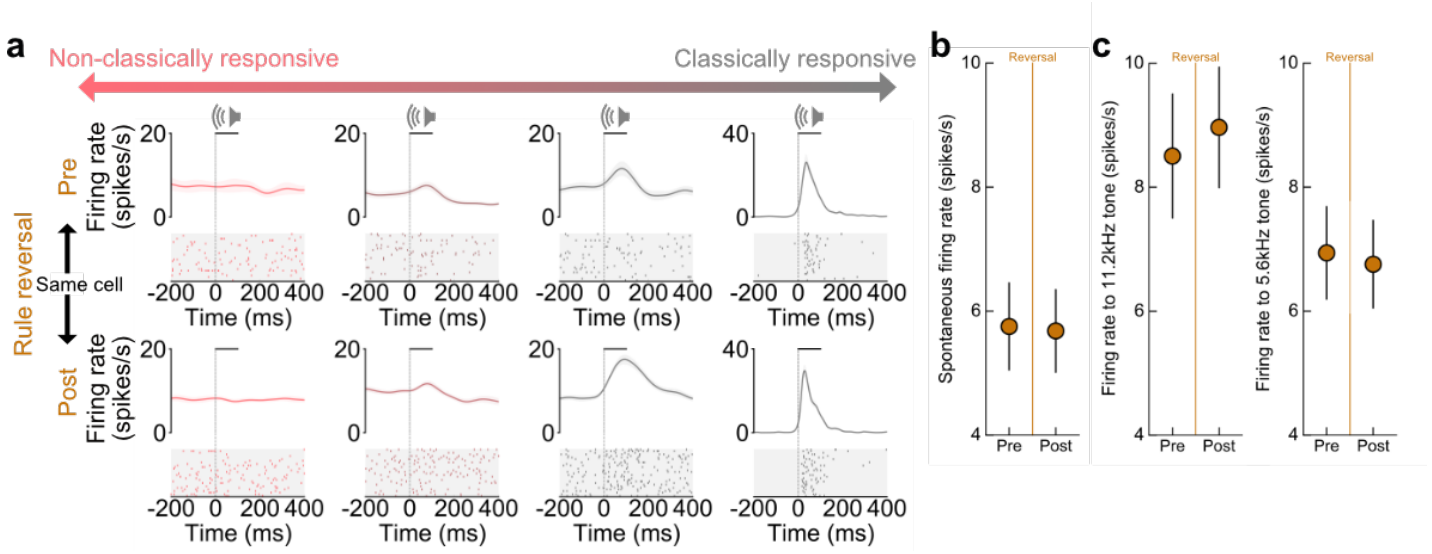


Extended Data Figure 4. Classically and non-classically responsive cells are similarly modulated by movement during behavior. a, Schematic of movement tracking pipeline. Mouth, nose, and whisker represent tracked key points. **b**, Average measured speed of tracked key points before and after the onset of movement. **c**, Average firing rates for classically responsive cells (n = 10 cells) and non-classically responsive cells (n = 7 cells) before and after the onset of movement. Lines, mean firing rate; shading, s.e.m. **d**. Statistical comparison of cells from **c** before and after the onset of movement. Only mouth movements moderately altered the firing rates of both CRs and NCRs. Mouth CR, $p = 0.01$; Mouth NCR, $p = 0.04$; two-sided Wilcoxon signed rank test with Benjamini-Hochberg corrections). Data are mean+s.e.m. *$p<0.05$.

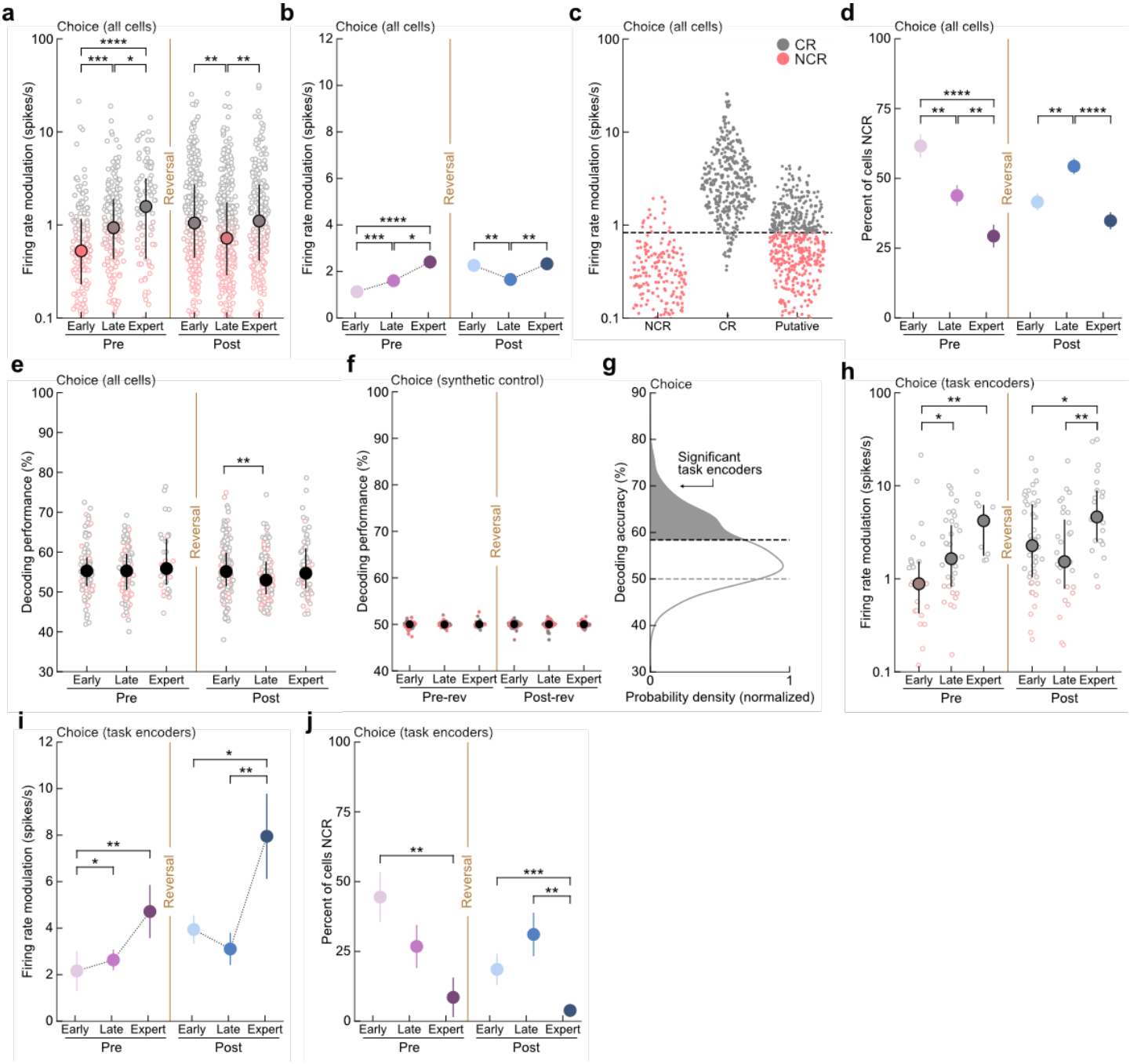


Extended Data Figure 5. Non-classically responsive auditory cortical neurons preferentially recruited during learning at single-cell and population level. a, Distribution of statistically defined classically and non-classically responsive neurons with respect to stimulus responses, as well as putative neurons. Dotted line is a separating threshold fitted using a support vector machine (described in Methods as "Discrete characterization of classically and non-classically responsive units"). **b**, Percent of stimulus classically responsive cells (including enhanced or suppressed cells) and non-classically responsive cells over reversal learning for all recorded cells. The percentage of non-classically responsive cells increases during the rapid learning phase (i.e. 'late') of both pre-and post-reversal. **c**, Single-cell, single-trial stimulus decoding performance relative to synthetic controls for the overall population during reversal learning. Pre_{expert} (n = 74 cells) vs. Pre_{early} (n = 113 cells), $p = 2x10^{-6}$; vs. Pre_{late} (n = 158 cells), $p = 6x10^{-7}$; $Post_{early}$ (n = 256 cells) vs. $Post_{late}$ (n = 299 cells), $p = 0.03$; vs. $Post_{expert}$ (n = 143 cells), $p = 3x10^{-5}$; $Post_{late}$ vs. $Post_{expert}$, $p = 0.003$. Data are median+i.q.r. **d**, Decoding performance from cells in **c** applied to spike trains synthetically generated by sampling (with replacement) from all observed ISIs; $p > 0.05$. Data are median+i.q.r. **e**, Left, mean stimulus firing rate modulation for all cells (overall population) from **Fig. 3a** and passive exposure cells from **Fig. 3b**. We observed lower firing rate modulations during the late phases of both pre-and-post reversal and $Post_{late}$ compared to passive controls (Pre_{late}, $p = 0.04$; $Post_{late}$, $p = 1x10^{-7}$; $Post_{expert}$, $p = 1x10^{-5}$). Right, percent of cells from the overall population that are stimulus non-classically responsive cells. We observed a significant increase in the percent of non-classically responsive cells during the late phases of both pre-and-post reversal and $Post_{late}$ relative to passive controls (NCR, Pre_{late}, $p = 0.03$; $Post_{late}$, $p = 4x10^{-6}$; $Post_{expert}$, $p = 2x10^{-5}$), bootstrapped hypothesis test with Benjamini-Hochberg corrections. **f**, Percent of

task-encoding (stimulus) classically responsive and non-classically responsive cells during reversal learning. The percentage of task-encoding non-classically responsive cells increased during the rapid learning phase (i.e. 'late') of both pre-and post-reversal. **g**, Ensemble stimulus decoding performance for ensembles of up to 20 cells from the overall population as a function of ensemble composition during pre-reversal. While early in learning ensembles composed of mostly classically responsive cells had the highest decoding performance, during late learning ensembles composed of mostly non-classically responsive cells had the highest decoding performance. Pre_{early} (n = 1,884 ensembles); Pre_{late} (n = 1,584 ensembles); Pre_{expert} (n = 1,259 ensembles); *p ≤ 0.05, **p ≤ 0.01, ***p ≤ 0.001, ****p ≤ 0.0001; all tests vs. 0-10% NCR ratio. **h**, Ensemble stimulus decoding performance for ensembles of up to 20 cells from the overall population as a function of ensemble composition during post-reversal. While early in learning ensembles composed of mostly classically responsive cells had the highest decoding performance, during late and expert phases ensembles composed of mixed ensembles (60% NCR) had the highest decoding performance. $Post_{early}$ (N = 3,081 ensembles); $Post_{late}$ (N = 3,342 ensembles); $Post_{expert}$ (N = 1,541 ensembles); *p ≤ 0.05, **p ≤ 0.01, ***p ≤ 0.001, ****p ≤ 0.0001; all tests vs. 0-10% NCR ratio. **i,** Percent of cells from **Fig. 3k** that are stimulus non-classically responsive (NCR) across pre-reversal, first reversal, and second reversal. We observed that progressively more NCRs are recruited as animals undergo multiple reversals. Pre_{expert} vs. $Second_{expert}$, $p = 4x10^{-6}$; $First_{expert}$ vs. $Second_{expert}$, p = 0.01; bootstrapped hypothesis test with Benjamini-Hochberg corrections. Data are mean+s.e.m unless otherwise stated. *p<0.05, **p<0.01, ***p<0.001, ****p<0.0001; two-sided Mann-Whitney U test with Benjamini-Hochberg corrections.



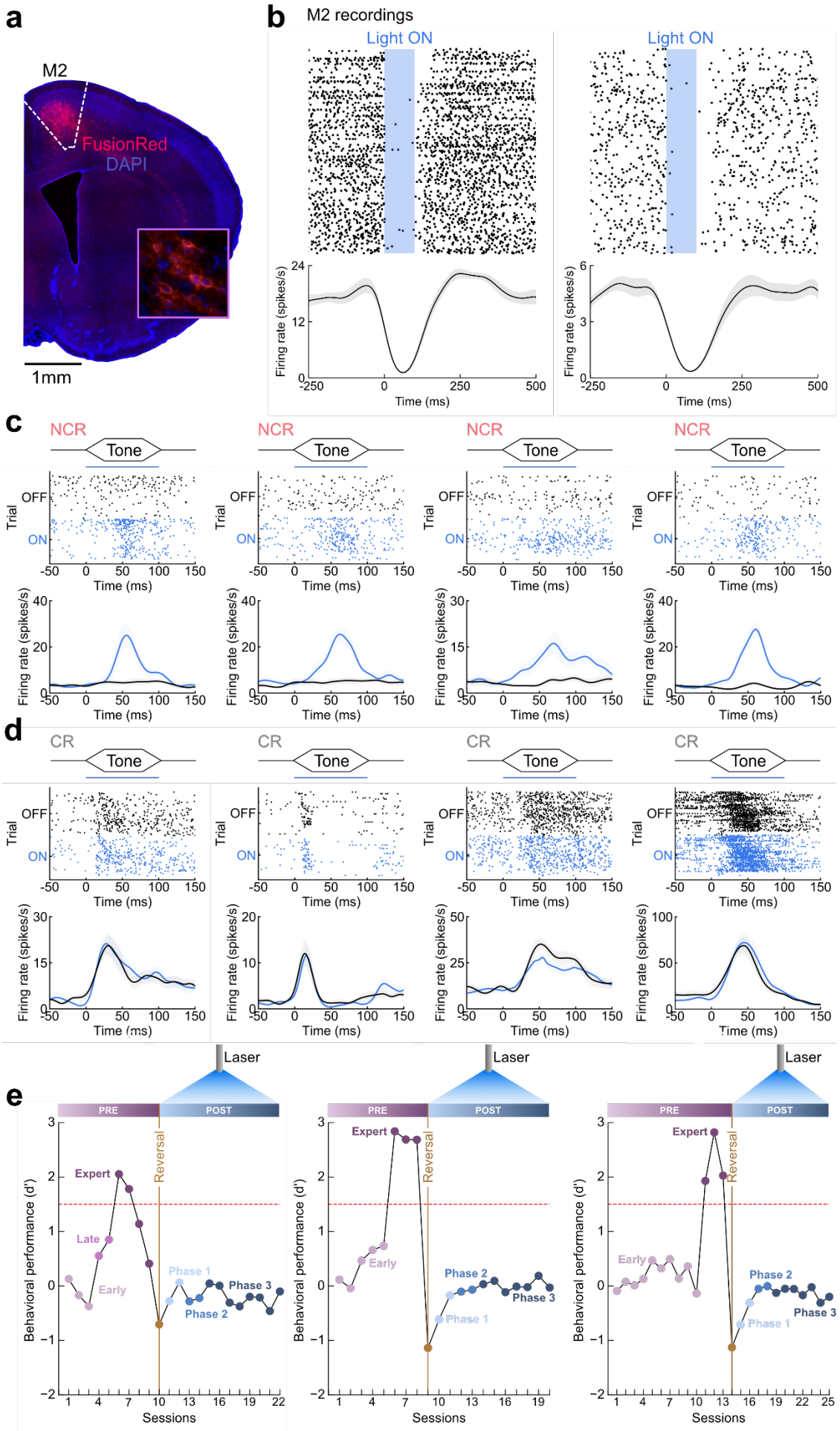
Extended Data Figure 6. Neural and behavioral responses during rule reversal session. a, Rasters and peri-stimulus time histograms (PSTHs) for 4 cortical neurons exemplifying the range of stimulus responses from non-classically responsive (red) to classically responsive (gray) during a rule reversal behavioral session. The same cells were recorded during pre and post rule reversal. Trials aligned to stimulus presentation. Horizontal bar, tone duration. Lines in PSTH, mean firing rate; shading, s.e.m. For rasters, a subset of trials is shown for clarity. **b**, Spontaneous firing rate before and after rule reversal; n = 154 cells, p = 0.488. **c**, Left, firing rate to the pre-reversal target tone (11.2kHz) before and after rule reversal, p = 0.6; Right, firing rate to the post-reversal target tone (5.6kHz), before and after rule reversal, p = 0.8. Data are mean+s.e.m.; two-sided Mann-Whitney U test.



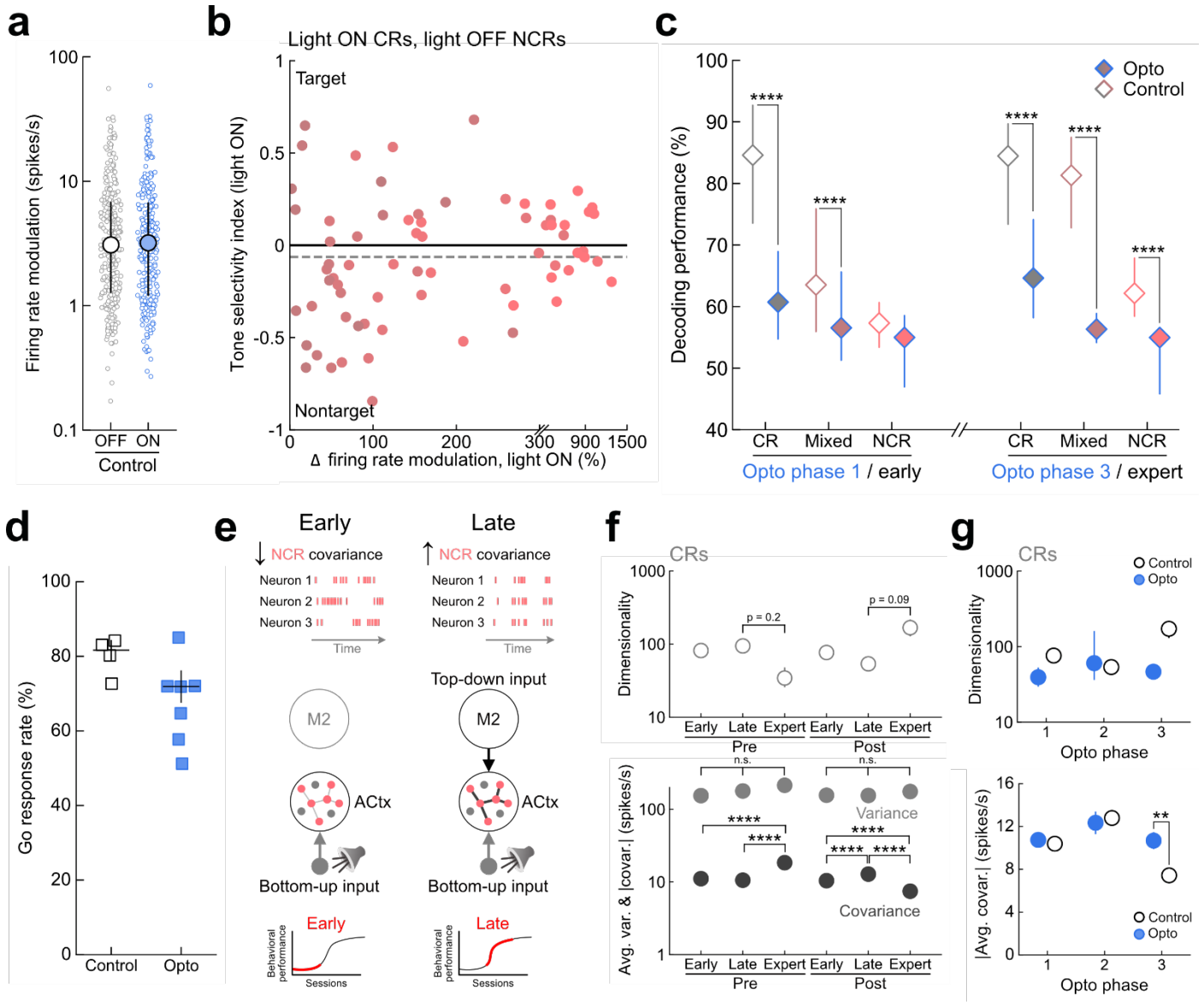
Extended Data Figure 7. Choice modulated cells during reversal learning.

a, For choice-related neural responses the firing rate modulations of the overall population monotonically increased over pre-reversal learning with a decrease during post-reversal late learning, although the effect size was marginal compared to stimulus-related activity. Distributions of choice firing rate modulations during reversal learning. Pre_{early} (n = 148) vs. Pre_{late} (n = 187), p = $2x10^{-4}$; vs. Pre_{expert} (n = 97), p = $5x10^{-6}$; Pre_{late} vs. Pre_{expert}, p = 0.01; $Post_{early}$ (n = 306) vs. $Post_{late}$ (n = 342), p = 0.001; $Post_{late}$ vs. $Post_{expert}$ (n = 220), p = 0.002. Data are median+i.q.r. **b**, Average choice firing rate modulation distributions of cells from **a** during reversal learning. **c,** Distribution of statistically defined choice classically and non-classically responsive neurons, as well as putative neurons. Dotted line is a separating threshold fitted using a support vector machine (described in Methods as "Discrete characterization of classically and non-classically responsive units"). **d**, Percent of choice non-classically responsive cells decreases during pre-reversal learning and transiently increases during post-reversal late learning in the overall population. Pre_{early} vs. Pre_{late}, p = 0.001; vs. Pre_{expert}, p = $4x10^{-6}$; Pre_{late} vs. Pre_{expert}, p = 0.007; $Post_{early}$ vs. $Post_{late}$, p = 0.001; $Post_{late}$ vs. $Post_{expert}$, p = $4x10^{-6}$; bootstrapped hypothesis test with Benjamini-Hochberg corrections. **e**, Single-cell, single-trial choice decoding performance across learning for cells from **a** relative to synthetic controls (n = 653). $Post_{early}$ vs. $Post_{late}$, p = 0.008. Data are median+i.q.r. **f**,

Choice decoding performance from the overall population of cells applied to spike trains synthetically generated by sampling (with replacement) from all observed ISIs, $p > 0.05$. Data are median+i.q.r. **g**, Choice decoding performance for all cells across all phases, highlighting cells with significant decoding performance (i.e. 'task-encoders'). N = 179/680 cells; test against one-sided 95% confidence interval set by below chance decoding neurons. **h**, The shift in distributions of choice firing rate modulations for task-encoding cells across learning was consistent to the overall population seen in **a**. Pre_{early} (n = 27 cells) vs. Pre_{late} (n = 37 cells), p = 0.03; vs. Pre_{expert} (n = 12 cells), p = 0.006; $Post_{expert}$ (n = 22 cells) vs. $Post_{early}$ (n = 49 cells), p = 0.02; vs. $Post_{late}$ (n = 32 cells), p = 0.006. **i**, Average choice firing rate modulation distributions of task-encoding cells from **h** during reversal learning. **j**, Percent of cells from **h** that are choice non-classically responsive. Pre_{early} vs. Pre_{expert}, p = 0.002; $Post_{expert}$ vs. $Post_{early}$, p = 0.0004; vs. $Post_{late}$, p = 0.001; bootstrapped hypothesis test with Benjamini-Hochberg corrections. Data are mean+s.e.m except where otherwise stated. *p<0.05, **p<0.01, ***p<0.001, ****p<0.0001; two-sided Mann-Whitney U test with Benjamini-Hochberg corrections


a
M2
FusionRed
DAPI
1mm
b
M2 recordings
Light ON
Light ON
Firing rate (spikes/s)
Time (ms)
c
NCR
Tone
Trial
OFF
ON
Time (ms)
Firing rate (spikes/s)
d
CR
e
Laser
PRE
POST
Reversal
Expert
Late
Early
Phase 1
Phase 2
Phase 3
Behavioral performance (d')
Sessions

Extended Data Figure 8. Photoinhibition of M2 inputs modulates NCRs and impairs reversal learning. a, Immunohistochemical confirmation of inhibitory stGtACR2 expression in secondary motor cortex. **b**, Example silicon-probe recordings from two M2 neurons containing stGtACR2. Optogenetic stimulation efficiently suppressed neuronal spiking for the duration of the laser pulse. Blue shaded area = light ON. Lines in PSTH, mean firing rate; shading, s.e.m. **c**, Non-classically responsive cells become classically responsive when M2 inputs to ACtx are inhibited. Example rasters and PSTHs from 4 neurons in ACtx that are non-classically responsive (NCR) during light OFF trials but demonstrated enhanced firing in the light ON condition. Horizontal bar in rasters, laser duration. Lines in PSTH, mean firing rate; shading, s.e.m. **d**, Classically responsive cells are unaltered by photoinhibition of M2 inputs. Example rasters and PSTHs from 4 neurons in ACtx that are classically responsive (CR) during light OFF trials and demonstrated no change in firing during the light ON condition. Lines in PSTH, mean firing rate; shading, s.e.m. **e**, Individual learning curves from 3 photoinhibited animals during post-reversal. Each panel shows behavioral d' values across all reversal learning behavioral sessions. Top schematic shows duration of optogenetic inactivation. Early, late, and expert phases of pre-reversal and opto phases 1, 2, and 3 are indicated on each learning curve.



Extended Data Figure 9. Photoinhibition of M2 inputs during learning suppresses both target and nontarget-tuned cells and impairs stimulus decoding performance. a, Average stimulus firing rate modulation of all neurons recorded in ACtx of sham control in light OFF and ON conditions (n = 267 cells), p = 0.2, two-sided Wilcoxon signed-rank test. **b**, Tone selectivity index for cells that are non-classically responsive (NCRs) in the light OFF condition but become tone responsive (CRs) in the light ON condition (n = 72 cells, values between 0 and 1 = tuning to target tone, values between 0 and -1 = tuning to non-target tone, 0 = equal modulation by both tones).Tone selectivity is assessed during light ON conditions, p = 0.1; two-sided one-sample t-test. **c**, Photoinhibition impairs ensemble decoding performance for all ensemble types. Left, ensemble decoding performance during Opto phase 1 (the equivalent of $Post_{early}$). Classically-responsive ensembles (Opto, n = 2,016 ensembles vs. control, n = 1,462 ensembles, p = $1x10^{-292}$); Mixed ensembles comprised of both CRs and NCRs (Opto, n = 562 ensembles vs. control, n = 963 ensembles, p = $1x10^{-25}$); Non-classically responsive ensembles (Opto, n = 38 ensembles vs. control, 328 ensembles, p = 0.08). Right, ensemble decoding perfromance during Opto phase 3 (the equivalent of $Post_{expert}$). Classically-responsive ensembles (Opto, n = 334 ensembles vs. control, n = 117 ensembles, p = $6x10^{-22}$); Mixed ensembles comprised of both CRs and NCRs (Opto, n = 244 ensembles vs. control, n = 659 ensembles, p = $1x10^{-94}$); Non-classically responsive ensembles (Opto, n = 16 ensembles vs. control, 171 ensembles, p = $7x10^{-7}$). Data are median+i.q.r. **d**, Photoinhibition of M2 inputs do not result in more Go responses. Response rates at start of post-reversal; opto vs. sham, p = 0.07. **e**, Left, early in learning, ACtx receives primarily bottom up sensory input, and activity of non-classically responsive neurons in ACtx are decorrelated. Right, as learning progresses, M2 inputs recruit highly-correlated non-classically responsive activity

producing a new subspace that enables learning. **f**, Top, dimensionality of activity from all classically responsive neurons in ACtx over learning. Dimensionality decreases during the transition from late to expert in both phases. Pre_{late} vs. Pre_{expert}, $p = 0.2$; $Post_{late}$ vs. $Post_{expert}$, $p = 0.09$; bootstrapped hypothesis test with Benjamini-Hochberg corrections. Data are median+i.q.r. Bottom, average variance and absolute covariance between all classically responsive neurons in ACtx over learning. Light gray, variance, $p > 0.05$. Dark gray, absolute covariance, Pre_{late} vs. Pre_{expert}, $p = 2x10^{-7}$; Pre_{early} vs. Pre_{expert}, $p = 2x10^{-7}$; $Post_{early}$ vs. $Post_{late}$, $p = 8x10^{-5}$; $Post_{late}$ vs. $Post_{expert}$, $p = 2x10^{-9}$; $Post_{early}$ vs. $Post_{expert}$, $p = 1x10^{-5}$; two-sided student's t test with Benjamini-Hochberg corrections. **g**, Top, dimensionality of activity of all classically responsive neurons in ACtx in photoinhibited vs. control animals over post-reversal, $p > 0.05$, bootstrapped hypothesis test with Benjamini-Hochberg corrections. median+i.q.r. Bottom, average absolute covariance between all classically responsive neurons in ACtx in photoinhibited vs. control animals over post-reversal. Phase 3 vs. $Post_{expert}$, $p = 0.001$, two-sided student's t test with Benjamini-Hochberg corrections. Data are mean+s.e.m unless otherwise stated. **p<0.01, ***p<0.001, ****p<0.0001; two-sided Mann-Whitney U test with Benjamini-Hochberg corrections.