Prediction error signals in anterior cingulate cortex drive task-switching

Nicholas Cole[1], Matthew Harvey[1], Dylan Myers-Joseph[1], Aditya Gilra[2,3], Adil G. Khan[1]

1: Centre for Developmental Neurobiology, King's College London, London, UK

2: Machine Learning Group, Centrum Wiskunde & Informatica, Amsterdam, The Netherlands

3: Department of Computer Science, The University of Sheffield, Sheffield, UK

Correspondence to Adil G. Khan: khan.adil@kcl.ac.uk

Abstract

Task-switching is a fundamental cognitive ability which requires animals to update their knowledge of current rules, allowing flexible behaviour in a changing environment[1]. This is often achieved through evaluating discrepancies between observed and expected events. The anterior cingulate cortex (ACC) has a key role in processing such discrepancies, or prediction errors[2–12]. However, the neural circuit mechanisms underlying task-switching are largely unknown. Here we show that activity in the ACC induced by the absence of expected stimuli is necessary for rapid task-switching. Mice trained to perform a block-wise set-shifting task typically required a single experience of an expectation violation, or prediction error to accurately switch between responding to the same stimuli using distinct rules. Neurons in the ACC explicitly represented these prediction errors, and their activity was predictive of successful one-shot behavioural transitions. Prediction error signals were projection target-specific, constrained in their spatio-temporal spread across cortex, and heavily disrupted by VIP interneuron perturbation. Optogenetic silencing and single-trial un-silencing revealed that the requirement of the ACC in task-switching was restricted to the epochs when neural prediction error signals were observed. These results reveal a dedicated circuitry promoting the transition between distinct cognitive states.

Main

Cognition involves transitioning between multiple mental states encoding distinct rules to guide behaviour[13–15]. The ACC plays a central role in detecting the need for updating such behavioural rules, and implementing the updated rules[2–12,16–19]. While there is evidence for prediction error signalling in the prefrontal cortex (PFC) in cognitive tasks across humans, monkeys and rodents[2,3,10,11,20–22], it is unclear how such cognitive prediction error circuits are organised, what local inhibitory circuitry underlies their function, and to what extent prediction error responses have any causal influence on the subsequent updating of the animal's behavioural strategy.

To study the neural circuit basis of rapid task-switching, we trained head-fixed mice to switch between blocks of discriminating two visual stimuli or discriminating two olfactory stimuli (Fig. 1a, b). During olfactory discrimination, a random subset (70%) of odour stimuli were preceded by the same visual stimuli, now irrelevant to the task. Mice repeatedly switched between attending to and accurately discriminating the visual stimuli (rule 1), and ignoring the same visual stimuli while accurately discriminating the odour stimuli (rule 2). Mice

performed up to 15 behavioural switches, or block transitions per session (Fig. 1c, range 6 to 15, median 14 transitions across 13 sessions, 10 mice, one session per mouse shown. Number of trials per block, median ± IQR, 33±1 and 35±8 trials in visual and odour blocks respectively).

Our aim was to repeatedly capture the transition between two distinct and accurately applied task rules within a well-defined time period. Furthermore, to better separate cognitive processes from stimulus- or reward-evoked activity during the transition, we needed the block transitions to be inferred without any explicit stimulus or reward signal. Our task satisfied these requirements: on transitions from the odour to visual blocks, mice noticed the absence of an expected odour stimulus to switch their behaviour, and accurately responded to the now-relevant visual stimuli from the next trial onwards (Fig. 1d).

Optimal 'one-shot' transitions would require only a single error trial to successfully switch behaviour[23,24], as seen in Fig. 1d. Most odour to visual block transitions in our task were indeed one-shot (Fig. 1e, left, 56.9% one-shot transitions). In contrast to odour to visual block transitions inferred by the absence of odours, the visual to odour block transitions were characterised by the unexpected appearance of odour stimuli which had been absent during the visual block. These visual to odour block transitions were also largely one-shot (Fig. 1e, right, 54.4% one-shot transitions). The asymmetrical nature of the transitions allowed us to contrast inferred and explicitly cued behavioural transitions in the same task.

To understand the processes underlying this rapid task-switching, we first fit the behaviour to a basic reinforcement learning (RL) model (tabular SARSA, Fig. 1f, top) where an agent updated its estimated value of licking in response to stimuli in each trial (the Q-value). This model was unable to reproduce the rapid block transitions observed in the data (Fig. 1g top), and the best fit of learning and exploration rates produced slower block transitions (Fig. 1g middle, see methods). We next built an RL model with a context belief state, hierarchically controlling subsequent choices (Fig. 1f, bottom, context in this model refers to the different block rules). In this model the agent updated its belief about which context (or block) it was in by a context error signal. This model learnt two Q-value-tables for the distinct contexts, and subsequently switched between these two Q-value-tables on accumulating a large enough context error signal. This belief state model reproduced the rapid block transitions (Fig. 1g bottom), suggesting that rapid task-switching behaviour involves transitioning between abstract representations of rules or context, driven by prediction errors.

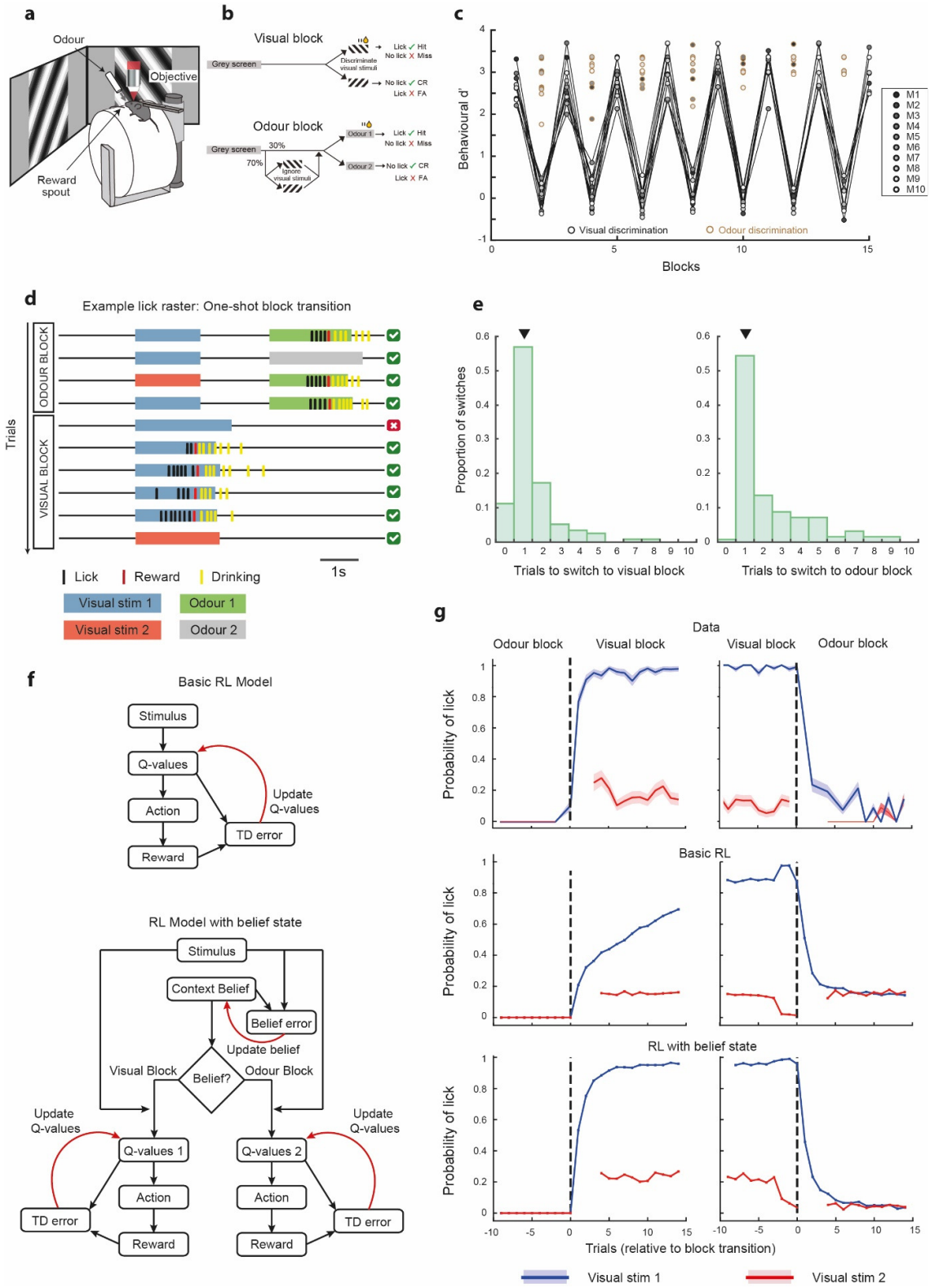


Figure 1 | Mice perform rapid task-switching with one-shot block transitions. a. Experimental setup. **b**. Task schematic. Mice switched between blocks of discriminating two visual stimuli in the visual blocks or discriminating two olfactory stimuli while ignoring the same visual stimuli in the odour blocks. **c.** Behavioural discrimination performance (behavioural d') across blocks (N = 10 mice). Shades

of grey indicate individual mice. Odour discrimination performance is shown in brown circles. Mice performed up to 15 block transitions in a session. **d.** Example behavioural data from an odour to visual block transition (lick raster) showing stimuli, lick and reward times. In the odour block the mouse ignored both visual gratings while accurately discriminating odour stimuli, but switched rules after a single error trial and accurately started discriminating the same visual stimuli. **e.** Histograms of the number of trials required to switch between the two blocks, with one-shot transitions indicated by arrowheads. Left, switch from odour to visual block, right, switch from visual to odour block. N = 85 and 93 transitions respectively, 13 sessions, 10 mice. **f**. Schematic of reinforcement learning (RL) models. Top, basic RL, bottom, RL with belief state. **g**. Average probability of licking the reward spout in response to Visual Stimulus 1 (rewarded only in visual blocks) and Visual Stimulus 2 (unrewarded in both blocks), aligned to the block transitions. Top, data from average of 85 odour to visual and 93 visual to odour block transitions. Shading indicates SEM. First 3 trials of each block were forced to be Visual Stimulus 1. Middle, basic RL model fit to data. Bottom, RL model with belief state fit to data.

A number of studies have linked ACC activity to surprising events[10,25], errors[11,18] or negative feedback[26,27], which promote updating of behavioural policies[6,16], or promote control[5]. Thus, to study the neural basis of rapid task-switching, we recorded the activity of populations of layer 5 ACC neurons during this behaviour using chronic two-photon calcium imaging through a microprism (Fig. 2a). As expected, we found that individual ACC neurons responded to many task-related variables. (Fig. 2b). Subsets of cells responded to visual and odour stimuli, locomotion onsets, reward delivery and licking (Fig. 2b, Extended Data Fig. 2a). A binary classifier was able to accurately decode the identity of visual and odour stimuli after stimulus onsets from ACC neural population activity, and was further able to decode the block type both before and after stimulus onsets (Fig. 2c). Thus, ACC contained diverse signals relevant to the ongoing task.

To account for rapid one-shot task-switching, we expected to find a signal in the brain triggered by the prediction error event at the odour to visual block transitions: an absence of the expected odour. If such a signal were specific to the prediction error, this signal would be (1) absent in the final trials of the odour block when the odour was expected and actually delivered, (2) present when the odour was expected but absent, and (3) absent once again when the odour was no longer expected later in the visual block (schematic of the three conditions shown in Fig. 2d). We found cells in the ACC which satisfied these criteria, which we termed 'mismatch neurons'. Fig. 2e shows an example positive mismatch neuron which was suppressed in trials when the expected odour did arrive, but responded strongly during the odour prediction mismatch (when an odour was expected but did not arrive), and was not responsive later when an odour was no longer expected. The percentage of all recorded cells which showed any mismatch responses (see methods) was 9%. While we focused on positive mismatch responsive cells in this study (Fig. 2e, 27% of mismatch cells), we found a similar number of cells with negative mismatch responses (30% of mismatch cells, Extended Data Fig. 2b) and other combinations of activity profiles. Response inhibition by the arrival of the expected odour occurred for both rewarded and non-rewarded odours, suggesting that these cells did not reflect a simple negative reward signal (Extended Data Fig. 2c). Changes in running and licking could not account for the identification of mismatch cells (Extended Data Fig. 2e, f).

Mismatch neurons were not widespread throughout the brain, as they were virtually absent when we performed the same experiment while recording from the primary visual cortex (V1) (Fig. 2g, see also Fig. 3).

Do mismatch neurons broadcast their activity widely across the brain, or are they enriched or excluded from populations based on their projection target? We tested this in the same mice by selectively labelling the sub-population of ACC neurons which projected to the striatum, a major projection target of the PFC[28]. CTB-Alexa647 injections in the striatum identified striatal-projection neurons in the ACC (Fig. 2h). Although this population of neurons had mostly overlapping response properties with non-striatal projection neurons (Extended Data Fig. 2d), they had a significantly smaller proportion of mismatch neurons (Fig. 2h, Chi-squared test of proportion P = 0.0015, striatal projecting mismatch cells = 21/410, 5%, non-striatal projecting mismatch cells = 463/4888, 9%). This lower proportion of striatal projecting mismatch cells was outside the 99% confidence intervals of the non-striatal projecting mismatch cells (Fig. 2i, bootstrap test). Thus, mismatch responses in the ACC are not indiscriminately broadcast, but are organised in a projection target-specific manner.

We next asked how sustained the mismatch response was, within and across trials. Within a trial, the positive mismatch response peaked soon after the time of expected odour onset, but remained significantly higher than both its baseline and the actual odour response till the beginning of the next trial (Fig. 2j, Wilcoxon signed-rank test mismatch response vs baseline aligned to mismatch and next stimulus $P = 2.98 \times 10^{-15}$, $P = 3.34 \times 10^{-8}$ respectively, mismatch response vs odour response, $P=1.69 \times 10^{-28}$, $P = 5.45 \times 10^{-15}$ respectively). Across trials, the mismatch response was present only on the first mismatch trial, and rapidly decayed to non-significant amplitudes on subsequent trials (Fig. 2k), revealing the highly transient nature of this prediction error signal. These results provide direct evidence for predictive processing in the ACC during task-switching[29,30], and are consistent with a prominent theory suggesting that the PFC largely signals the non-occurrence of expected outcomes[31].

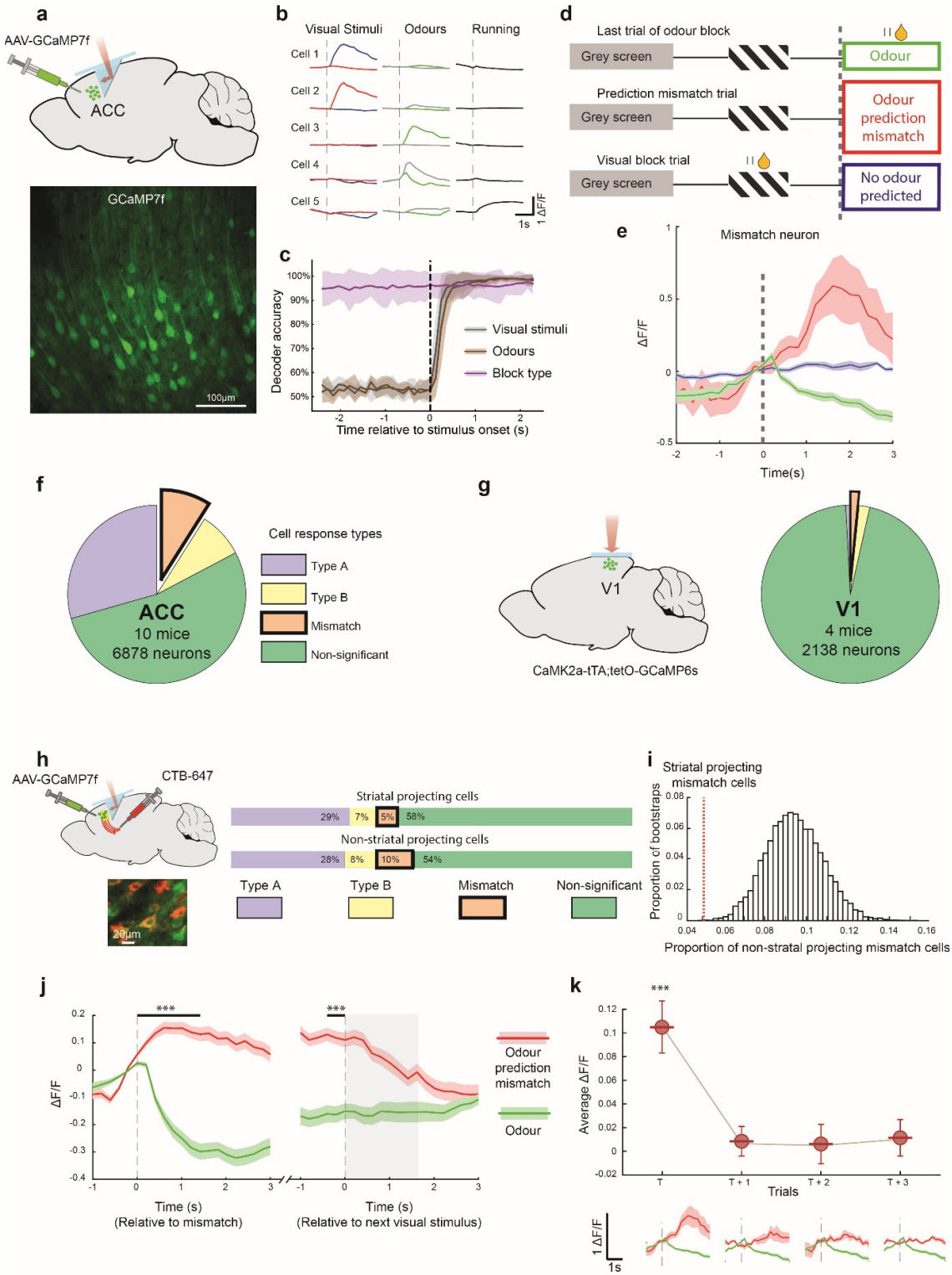


Figure 2 | ACC neurons encode prediction error signals. a. Top, schematic depicting two-photon imaging of neurons expressing GCaMP7f in the ACC through a microprism implant. Bottom, example ACC imaging site. **b.** Responses from 5 example neurons to task stimuli and running onsets. Shading here and below indicates SEM. **c**. Time course of decoding accuracy of a binary classifier using neuronal activity of simultaneously imaged cells from the ACC, mean of 9 sessions from 8 mice. Grey, decoding

the two visual stimuli in the visual block, brown, decoding the two odour stimuli in the odour block, and purple, decoding block-type from activity aligned to the non-rewarded visual stimulus onset in the two blocks. **d**. Schematic of behaviour during a one-shot transition from an odour to visual block. The first trial of the visual block contains an odour prediction mismatch, due to the absence of an expected odour stimulus, which mice can use to infer the change in block type. Mismatch cells were defined as neurons with significantly different activity between the odour prediction mismatch period and both the odour period and no odour predicted period. **e**. Example mismatch cell with a strong positive response to the odour prediction mismatch condition (red), compared to the actual delivery of the odour (green) or trials where odours were neither predicted nor delivered (blue). **f.** Proportions of cells that had significantly different activity (averaged 0 to 1.5s) between the three trial types described in **d**. Type A and Type B were significantly different only when comparing odour prediction mismatch to odour (N=2007 cells, 29%) or to no prediction conditions (N = 567 cells, 8%) respectively. Mismatch cells were significantly different in both comparisons (N = 616 cells, 9%, total 6878 cells, 10 mice). **g.** Left, schematic of two-photon calcium imaging from primary visual cortex (V1). Right, same comparisons as in e with cells recorded from V1 (N = 18 Type A, 1%, 41 Type B, 2%, and 35 mismatch cells, 1.6%, total 2138 cells, 4 mice) **h**. Left, schematic of cell labelling and example images of retrogradely labelled striatal-projecting (CTB-Alexa647 labelled) and non-striatal projecting cells in the ACC expressing GCaMP7f. Right, proportion of mismatch cells from striatal-projecting and non-striatal projecting neurons in ACC. Striatal projecting: N =124 Type A (29%), 30 Type B (7%), 21 mismatch cells (5%), total 421 cells. Non-striatal projecting: N = 1384 Type A (28%), 392 Type B (8%), 484 mismatch cells (10%), total 4888 cells, 8 mice). **i**. Bootstrapped distribution of proportion of mismatch cells in non-striatal projecting cells. Data for proportion of striatal projecting mismatch cells indicated as vertical line. Data is outside the 99% confidence intervals. **j.** Mean response of all positive mismatch neurons (N = 168 cells) aligned to the odour prediction mismatch event, and to the next visual stimulus onset (grey shading). Response shown to the odour prediction mismatch event (red) and actual odour delivery (green). Wilcoxon signed-rank test between the two conditions averaged 0 to 1.5s from mismatch event and -0.5 to 0 s from the next stimulus onset, *** $P = 1.69 \times 10^{-28}$ and $P = 5.45 \times 10^{-15}$ respectively. **k**. Mean odour prediction mismatch response relative to pre-event baseline aligned to the same temporal epoch over consecutive trials (T is first mismatch event in an odour to visual block transition), shown up to 3 trials following first mismatch event (Wilcoxon signed-rank test, average -0.5 to 0s compared to 0 to 1.5 s, *** $P = 2.98 \times 10^{-15}$). Bottom, example cell responses showing the rapid decay of mismatch response over trials.

Although we had established the near-absence of mismatch cells in V1 (Fig. 2g), it was unclear how widely the prediction-error signal extended across the cortex in this task. We studied the spatio-temporal spread of prediction-error signals across the entire dorsal cortical surface of the mouse brain using widefield calcium imaging (Fig. 3a). We aligned the hemodynamic-corrected calcium activity to the three temporal epochs of odour onset, odour prediction mismatch, and no odour predicted, as described in Fig. 2d. The activity in these epochs evolved across the cortex over time, with higher activity apparent in the odour prediction mismatch condition (Fig. 3b, first three columns). We found pixels at each time point which satisfied the same criteria as mismatch neurons (Fig. 2d-e, see methods) and mapped their activity (Fig. 3b final column, colour map represents t-statistic of comparison between first and second column). This revealed the spatio-temporal evolution of the prediction-error signal across dorsal cortex, which spread from anterior secondary motor regions to posterior areas to include posterior parietal cortex (average map from 1.2 to 2.7s shown in Fig. 3c.

Averaged activity PSTHs of the medial secondary motor cortex, superior to the ACC shown in Fig. 3d). Together this revealed a spatio-temporally restricted map of prediction error signalling, particularly in areas superior to prefrontal cortex.

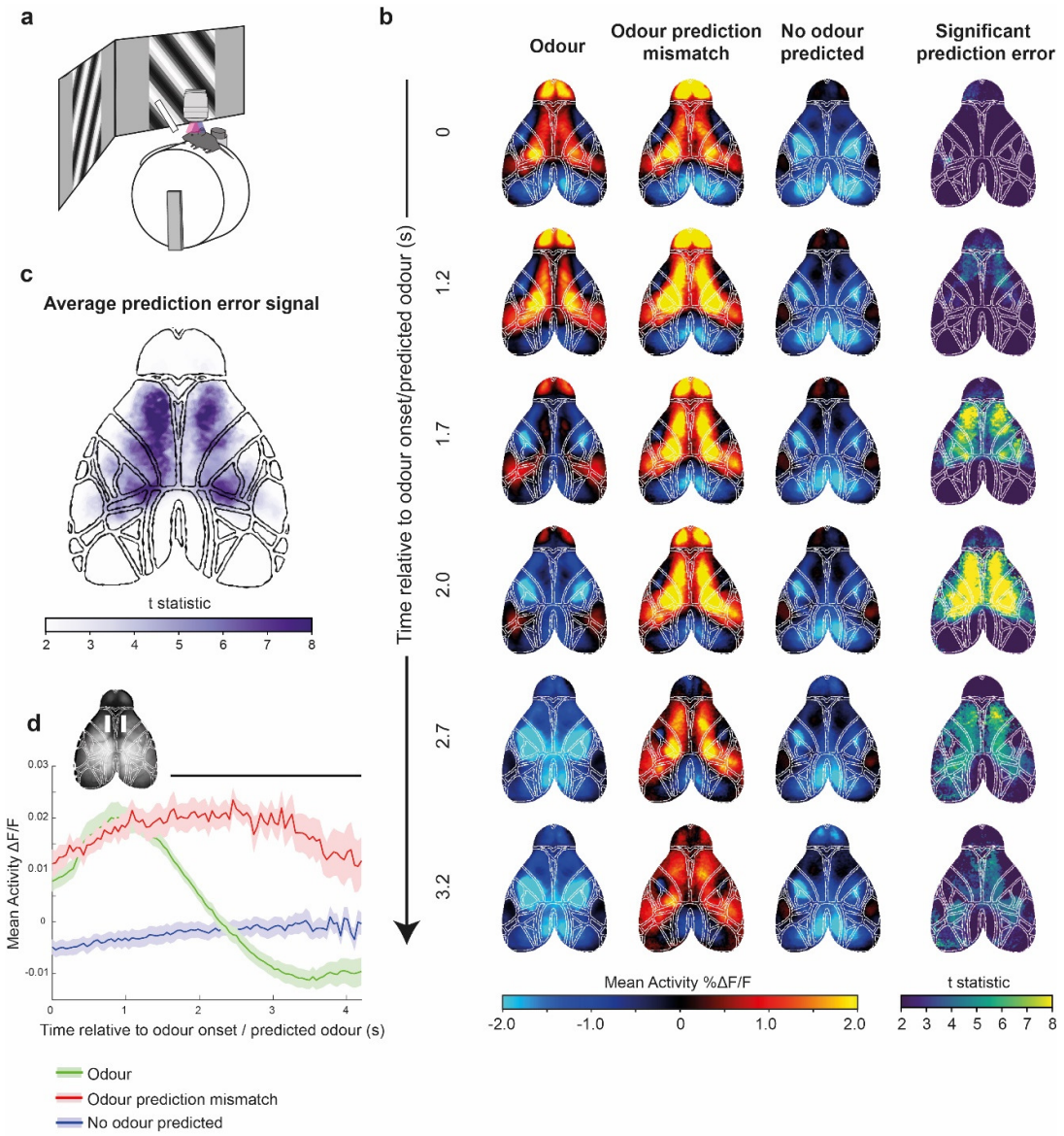


Figure 3 | Cortex-wide map of prediction error signalling. a. Schematic of widefield calcium imaging. **b.** Mean activity (columns 1-3) and significant prediction error pixels (column 4) across 6 timepoints, aligned to expected/actual odour onset (N = 11 sessions, 4 mice). Significance column shows pixels with activity significantly different between odour prediction mismatch and odour trials, and between odour prediction mismatch and no odour predicted trials, with the t-statistic values from the comparison of column 1 and 2 displayed in colour code. **c.** Pixels with significant prediction-error signalling, averaged across 1.2-2.7s relative to the expected/actual odour onset. **d.** Activity profile from the medial region of secondary motor cortex (inset). Periods during which the odour prediction mismatch response was significantly different from the odour and no odour conditions (paired t-test, $P < 0.01$, N = 11 sessions, 4 mice) shown as a black line above.

Local inhibitory circuits are crucial in shaping cortical activity[32], and vasoactive intestinal peptide (VIP) interneuron driven disinhibition through somatostatin (SOM) expressing interneurons is a cortical microcircuit motif[33–35] believed to be important in producing prediction-error responses in V1 during visuo-motor coupling mismatch[36]. We asked whether VIP interneurons played a role in producing mismatch responses in the ACC in our task. We employed an all-optical approach, where we photoactivated VIP interneurons while simultaneously measuring the activity of VIP and non-VIP neurons in the same ACC circuit with two-photon calcium imaging (Fig. 4a-c, 8/10 of the same mice as above). On photoactivating VIP cells in passive mice with increasing light powers (Fig. 4c inset), we observed monotonically increasing population activity (Fig. 4c), which demonstrated the effective disinhibition induced by our optogenetic manipulation. We photoactivated VIP cells during the task on each trial in the inter trial interval (ITI) period, which encompassed the prediction error event (Fig. 4d top). We compared the proportion of mismatch neurons in the VIP activation sessions to control sessions, and found that VIP activation during the ITI period strongly reduced the percentage of mismatch cells from 10% to 2% (Fig. 4d bottom, Chi-squared test of proportion $P < 0.0001$). We compared the average response amplitudes from positive mismatch neurons to the odour prediction mismatch period and odour period (Fig. 4e) between the VIP activated and control groups. We found a significant reduction only in the response to the odour prediction mismatch period during VIP cell activation, suggesting that VIP activation limits the mismatch response amplitude (Fig. 4e bottom). Critically, we confirmed that our VIP activation did not affect representations of stimuli and block rules in the ACC. A binary classifier as used in Fig. 2c was able to accurately decode the identity of visual and odour stimuli after stimulus onsets and the block type before and after stimulus onsets, both when VIP cells were photoactivated during the visual stimulus onsets (Fig. 4f) and during the ITI period (Fig. 4g). Thus, inducing disinhibition through VIP interneuron activation specifically and heavily disrupted prediction error signalling in the ACC.

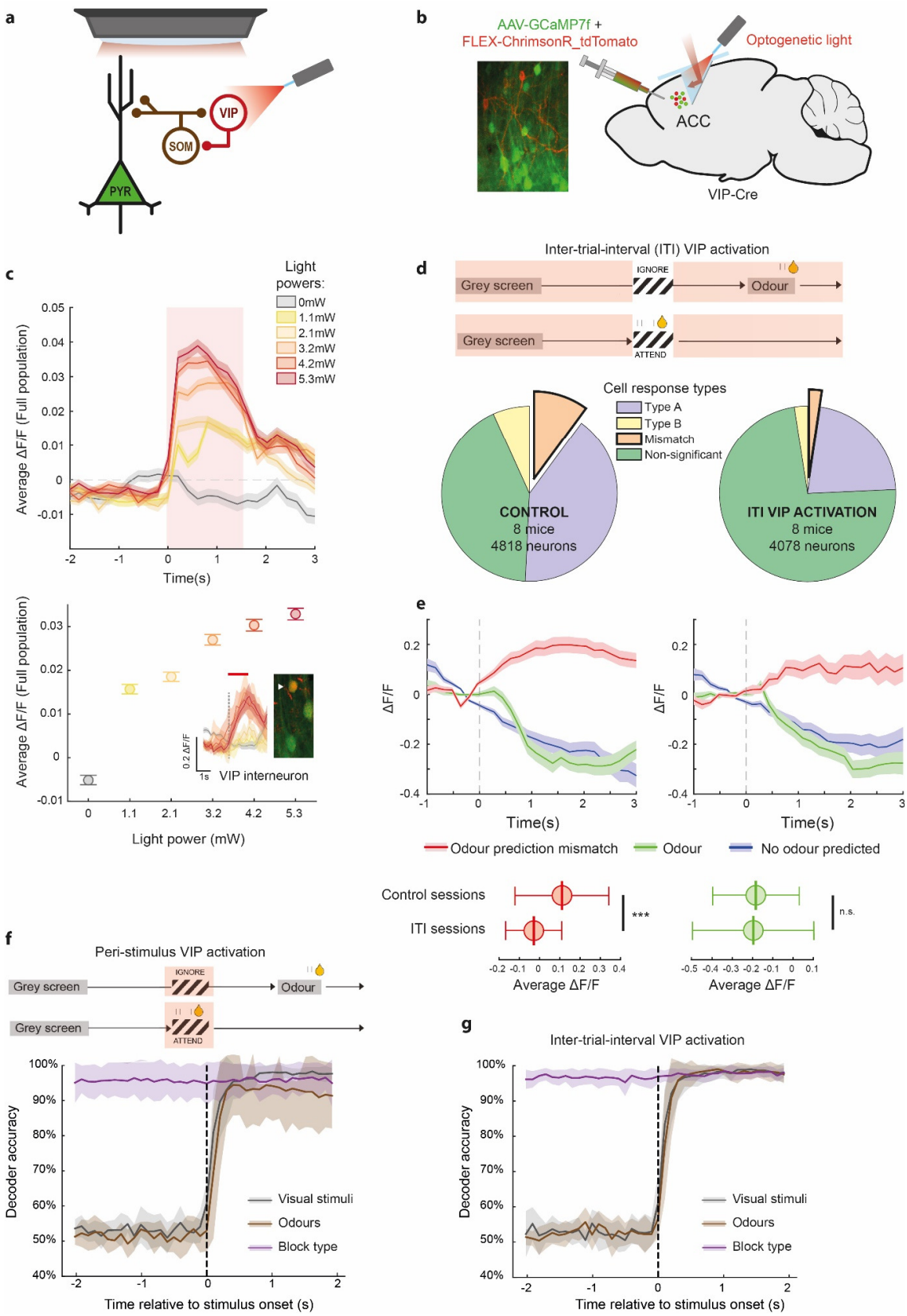

a
VIP
SOM
PYR
b
AAV-GCaMP7f +
FLEX-ChrimsonR_tdTomato
Optogenetic light
ACC
VIP-Cre
c
Light powers:
0mW
1.1mW
2.1mW
3.2mW
4.2mW
5.3mW
Average ΔF/F (Full population)
Time(s)
Light power (mW)
VIP interneuron
d
Inter-trial-interval (ITI) VIP activation
Grey screen
IGNORE
Odour
ATTEND
Cell response types
Type A
Type B
Mismatch
Non-significant
CONTROL
8 mice
4818 neurons
ITI VIP ACTIVATION
8 mice
4078 neurons
e
ΔF/F
Time(s)
Odour prediction mismatch
Odour
No odour predicted
Control sessions
ITI sessions
***
n.s.
Average ΔF/F
f
Peri-stimulus VIP activation
Grey screen
IGNORE
Odour
ATTEND
Decoder accuracy
Visual stimuli
Odours
Block type
Time relative to stimulus onset (s)
g
Inter-trial-interval VIP activation
Decoder accuracy
Visual stimuli
Odours
Block type
Time relative to stimulus onset (s)

Figure 4 | VIP interneuron photoactivation in ACC disrupts prediction-error signalling. a. Schematic of photoactivation of VIP cells to induce disinhibition of pyramidal neurons. **b.** Left, example site with neurons expressing GCaMP7f in green and VIP interneurons expressing the excitatory opsin Chrimson in red. Right, schematic of simultaneous 2-photon imaging and optogenetic activation of VIP interneurons in the ACC. **c.** Top, peri-stimulus time histograms of mean activity of all neurons (N = 2467 neurons, 8 mice) aligned to laser onset, at 5 different laser powers (maximum power 100% = 5.3mW), for 1.5s, mean ± SEM. Bottom, mean ± SEM activity of all neurons across laser window (0-1.5s). Inset: Image of an example VIP interneuron (arrowhead) and its response to increasing light powers, demonstrating photoactivation. **d.** Top, schematic of inter-trial-interval (ITI) activation epoch. Bottom, proportions of cells with significantly different activity between trial types described in Fig. 2d in control and ITI VIP activation sessions from the same sites, performed on subsequent days. **e**. Top, average activity in each trial type of positive mismatch cells identified from each session. Bottom, mean response (0 to 1.5s) of these cells to odour prediction mismatch (left, red) and odour (right, green). Wilcoxon rank-sum tests comparing control vs ITI sessions, mismatch responses $P = 1.1x10^{-4}$, odour responses P = 0.93. **f.** Top, schematic of peri-stimulus activation epoch. Bottom, time course of decoding accuracy of a binary classifier using neuronal activity during peri-stimulus VIP activation sessions (N = 4603 neurons, 8 mice), decoding task stimuli as in Fig. 2c. **g.** Time course of decoding accuracy with a binary classifier using activity from ITI VIP activation sessions, decoding the same stimuli.

To determine whether mismatch cells in the ACC actively contribute to the behavioural transitions across blocks, we asked if the amplitude of mismatch responses was predictive of success in rapid one-shot task-switching. We divided the odour to visual block transitions into one-shot and slower than one-shot behavioural transitions, or switches (Fig. 5a, N = 51% one-shot and 49% slower switches). Positive mismatch neurons were more active preceding one-shot, compared to slower switches (Fig. 5b) when aligned to the mismatch event (Wilcoxon signed-rank test, P = 0.0003) and when aligned to the next stimulus onset (P = 0.019, N = 146 cells). This result was also true when combining all mismatch neurons when aligned to the mismatch event ($P = 8.57x10^{-9}$, N = 616 cells).

Our belief-state RL model of the behaviour included an explicit context-prediction error signal, and we asked if this signal also predicted one-shot behavioural transitions across blocks. The model revealed a similar pattern, where the amplitude of the noisy prediction error signal was predictive of future one-shot transitions (Fig. 5c, Wilcoxon rank sum test, $P < 10^{-8}$, N = 70 transitions). These results support the idea that mismatch cells in the ACC play a key role in driving rapid task-switching behaviour, particularly when they produce larger prediction error signals.

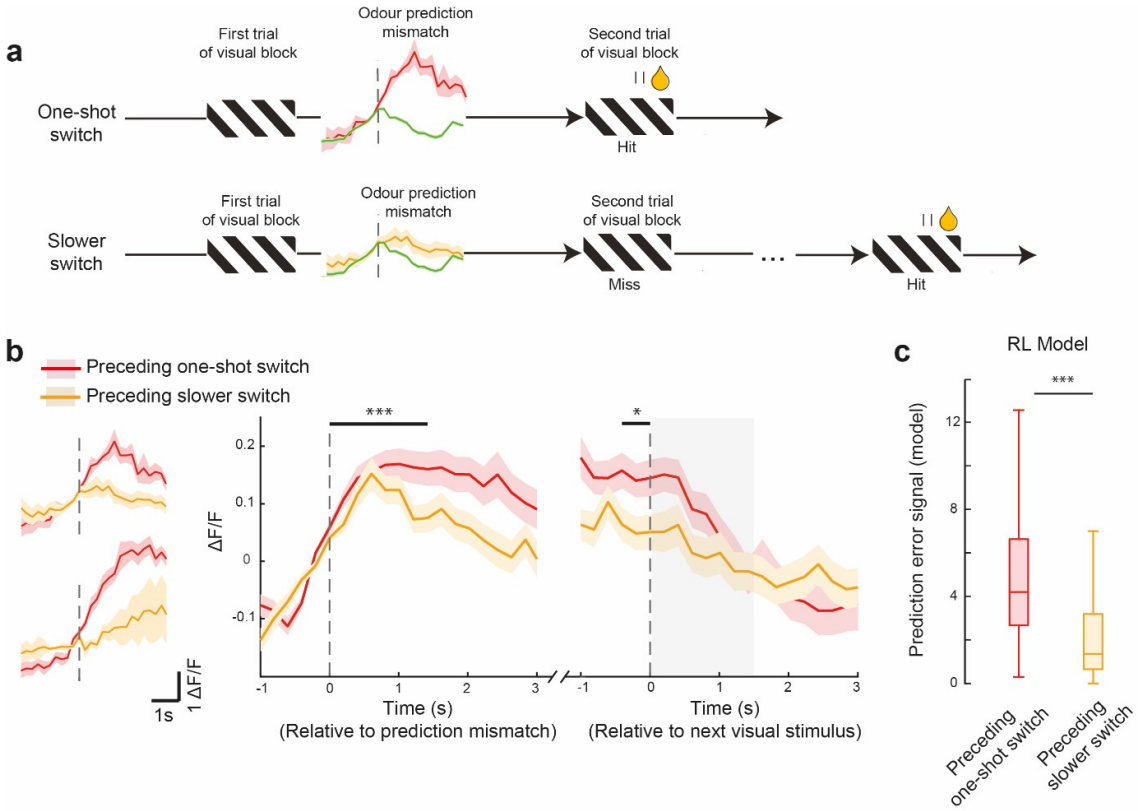


Figure 5 | Prediction error signals predict behaviour. a. Schematic showing odour to visual block transitions and the mismatch response preceding two scenarios, top, where the mouse executes a one-shot block transition, or switch, and bottom, where the switch is slower than one-shot. Red and yellow responses are mismatch responses from an example cell, green shows the response to delivery of odour, shading indicates SEM here and below. **b**. Left, odour prediction mismatch response of two example neurons preceding a one-shot (red) or slower switch (yellow). Right, mean responses across all positive mismatch neurons (N = 146 cells) preceding one-shot and slower switches (Wilcoxon signed rank test between the two conditions averaged 0 to 1.5 s from mismatch event and -0.5 to 0 s from the next stimulus onset, *** P = 0.0003, * P = 0.019). Grey shading indicates visual stimulus presentation. **c**. Prediction error signal amplitude from the Belief State RL model, preceding one-shot switches and slower switches, *** $P < 10^{-8}$, N = 70 transitions.

The opposite direction of block transition, from visual to odour blocks, was marked most saliently by the unexpected arrival of an odour. We again found mismatch cells in the ACC which responded differently to the unexpected odour, compared to the same odour when expected, or no odour (Extended data Fig. 3a, b, similar proportions of mismatch cells were obtained when controlling for running and licking behaviours, data not shown). Crucially, however, we obtained a similar proportion of these neurons in V1 (Extended data Fig. 3c), and furthermore, these mismatch neurons were not differently active preceding one-shot or slower block switches (similar analysis to Fig 5b, P > 0.05). Thus, mismatch cells signalling an unexpected appearance of an odour are found both within and beyond the ACC. Interestingly, neurons which represented both directions of mismatch were present at a higher proportion than expected by chance (Extended Data Fig. 3d), suggesting that prediction error signalling in the ACC may be a shared computation across distinct rule changes.

If prediction error signalling in the ACC is required for rapid task-switching, silencing the ACC would disrupt the behaviour specifically at block transitions. To test this, we optogenetically silenced the ACC bilaterally for entire sessions in a separate group of mice (Fig. 6a). ACC silencing caused a strong deficit in switching behaviour when the transition was inferred from the absence of an odour stimulus, that is, in the transition from odour to visual blocks. Strikingly, with the ACC silenced, mice ignored the rewarded visual stimulus repeatedly, consistent with the continuing application of the odour block rule (Fig. 6b). As a result, ACC silencing increased the number of trials taken to switch (Fig. 6c, d) and reduced the proportion of one-shot odour to visual block transitions (Fig. 6d, e). This was also the case when silencing ACC using bilateral infusions of the GABA-A receptor agonist muscimol (Extended Data Fig. 4a). In contrast, ACC silencing had no effect on the opposite direction of transition from visual to odour blocks, where the transition was marked by the arrival of a salient odour stimulus (Fig. 6d, Extended Data Fig. 4b). The belief state RL model could fit the ACC silencing data with the mismatch signal reduced by a factor of 0.22 in odour to visual block transitions, suggesting that the task-switching deficit could be explained by a partial reduction in the ACC mismatch signal (Extended data Fig. 4c). The model could also account for the insensitivity to ACC silencing in the visual to odour block transitions, with the mismatch signal reduced by a factor of 0.95 (Extended data Fig. 4d). Thus, in this task, the ACC is required for rapid task-switching, unless the appearance of highly salient stimuli overrides this requirement, possibly through comparable mismatch signals from other brain regions.

In addition to the ACC, the prelimbic cortex (PL) has been implicated in task-switching behaviour[37]. We optogenetically silenced PL bilaterally in the same mice for the entire session. PL silencing did not affect switching behaviour (Fig. 6c, e), indicating that the ACC has a specific role in updating behavioural rules, and this role is not widely shared with other prefrontal areas.

If the ACC indeed enables block transitions through prediction error signalling, the rapid decay of the mismatch signal over trials (Fig. 2k) would predict that the ACC is only required during the mismatch trial, and not earlier or later. To test this, we continuously silenced the ACC for the entire session and unsilenced it only on the first trial of a visual block when the prediction error occurs (Fig. 6f, top). The resulting behaviour showed no deficits in rapid switching (Fig. 6f, bottom, Extended data Fig. 4e), and the mice performed highly accurate discrimination for the rest of the block despite ACC silencing. Thus, the ACC is specifically required only for processing the prediction error and is no longer required for subsequent task performance. This conclusion is supported by the finding that during continuous ACC silencing, once the mice did eventually switch to discriminating the visual stimuli, the subsequent accuracy of visual discrimination was not significantly different from unsilenced sessions (Fig. 6g).

In addition to rapid task-switching, the ACC has been implicated in guiding slower learning processes[38–40]. We optogenetically silenced the ACC in a subset of mice as they first learned the switching task and found no effect of silencing on either the rate of learning a novel stimulus-reward association, or learning to ignore irrelevant stimuli[41] (Extended Data Fig. 4f). Thus, in our paradigm, the role of the ACC is specific to rapid task-switching and does not extend to guiding learning.

Finally, we asked whether the ACC was required at a specific temporal epoch within a trial, and whether this was consistent with the timecourse of the mismatch signal. We divided each trial into the inter trial interval (ITI) and the peri-stimulus time period (Fig. 6h). Silencing the

ACC both during the ITI and peri-stimulus period on each trial caused significant deficits in the switching behaviour, although each to a smaller degree than continuous silencing (Fig. 6h). Deficits in switching behaviour during both ITI and peristimulus ACC silencing were consistent with the timecourse of the mismatch signal, which extended through the ITI period up to the onset of the subsequent visual stimulus (Fig. 2i). Thus, ACC activity is required not only around the moment of expectation mismatch when the largest mismatch responses are present, but remains important until the next trial as the animal updates its belief about the current rule, and is thereafter no longer required for accurate behaviour.

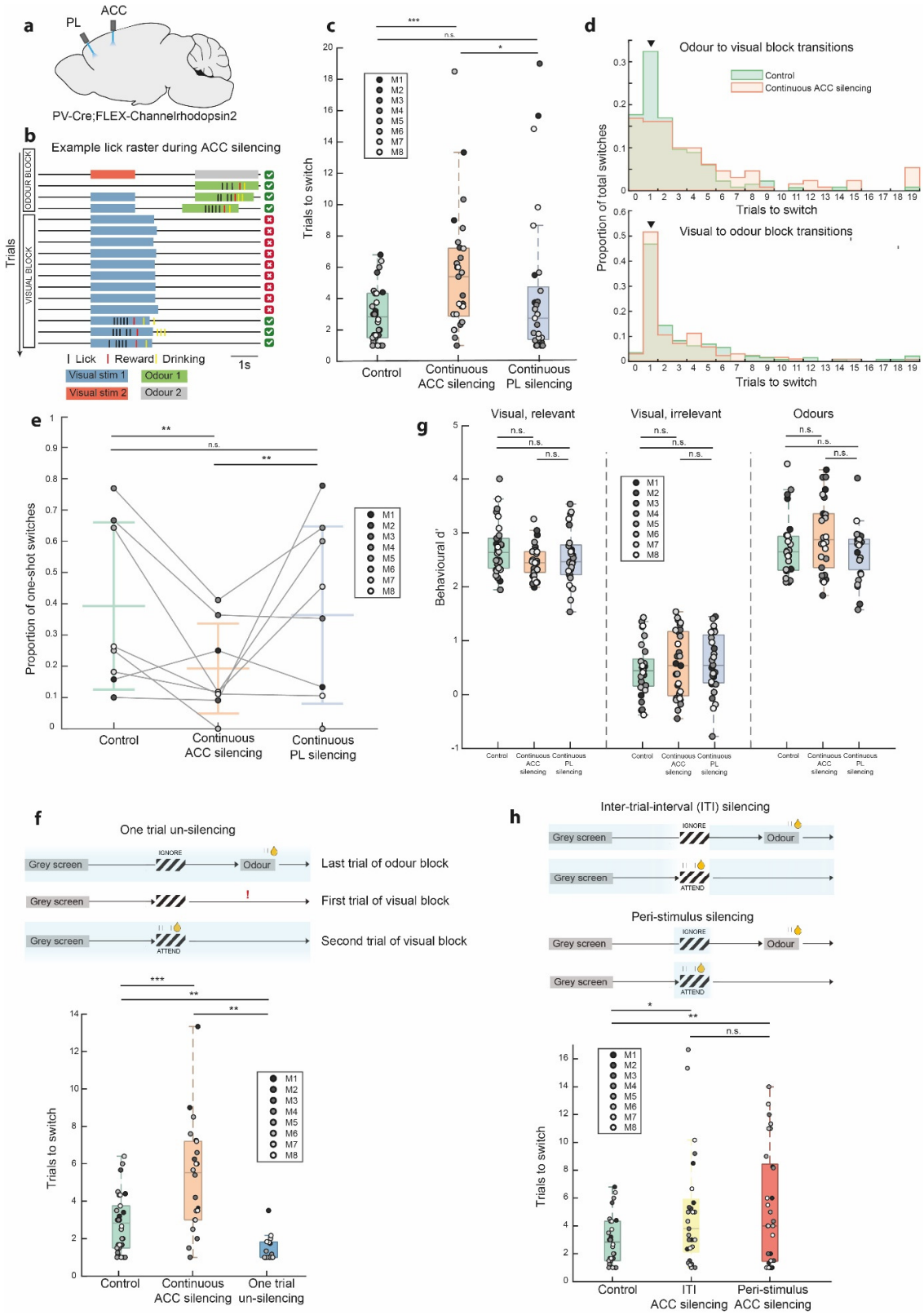


Figure 6 | Optogenetic inhibition of ACC impairs task-switching. a. Schematic of bilateral optic fibre implantation targeting both ACC and PL in each mouse. **b**. Example lick raster showing stimuli, lick and reward times during a transition from an odour to visual block during ACC silencing, showing impaired

switching behaviour. Compare to Fig. 1d. **c.** Average number of trials required in a session to switch from an odour to visual block, N = 30 sessions, 8 mice, 3-4 sessions per silencing condition per mouse, shades of grey indicate individual mice here and below. Median ± interquartile range (IQR) for control 2.83 ± 2.83 trials, continuous ACC silencing 5.4 ± 4.34 trials, continuous PL silencing 2.75 ± 3.37 trials. Wilcoxon signed-rank test comparing control and continuous ACC silencing P = 0.0006, control and continuous PL silencing P = 0.78, continuous ACC silencing and continuous PL silencing P = 0.02. **d.** Histogram of number of trials taken to switch between blocks during control and continuous ACC silencing sessions. Top, transitions from odour to visual block, N = 136 and 131 control and ACC silencing transitions respectively, bottom, transitions from visual to odour block, N = 148 and 146 control and ACC silencing transitions respectively. One-shot transitions indicated by arrowheads. **e.** Proportion of transitions from odour to visual block that are one-shot. Grey circles represent means for individual mice (N = 8), lines represent means ± STD across mice. Mean ±STD for control 0.39 ± 0.27, continuous ACC silencing 0.19 ± 0.14, continuous PL silencing 0.36 ± 0.28. Chi-squared test of proportions comparing control and continuous ACC silencing P = 0.001, control and continuous PL silencing P = 0.65, continuous ACC silencing and continuous PL silencing P = 0.005. **f.** Top, schematic depicting silencing during the entire session except for one trial at each odour to visual block transition. Red exclamation mark indicates odour prediction mismatch. Bottom, average number of trials required in a session to switch from an odour to visual block, median ± IQR here and below, control 2.83 ± 2.25 trials, continuous ACC silencing 5.4 ± 4.34 trials, one trial un-silencing 1.0 ± 0.82 trials, Wilcoxon signed rank test comparing control and continuous ACC silencing sessions P = 0.0006, control and one trial un-silencing P = 0.0084, continuous ACC silencing and one trial un-silencing P = 0.0016. **g.** Steady-state task discrimination performance after each successful block transition. Behavioural d-primes to visual stimuli in a visual block: control, continuous ACC silencing and continuous PL silencing sessions, 2.64 ± 0.55, 2.44 ± 0.38 and 2.46 ± 0.55 respectively, visual stimuli in an odour block 0.44 ±0.5, 0.54 ± 1.19, and 0.54 ± 0.89 respectively, and odour stimuli 2.64 ± 0.63, 2.87 ± 1.0, and 2.79 ± 0.57 respectively. Wilcoxon signed-rank tests as indicated, all Ps > 0.05. **h.** Top, schematics depicting inter-trial-interval (ITI) and peri-stimulus silencing epochs. Bottom, average number of trials required in a session to switch from an odour to visual block, control 2.83 ± 2.83 trials, ITI ACC silencing 3.8 ± 3.79 trials, and peri-stimulus ACC silencing 4.0 ± 6.98 trials. Wilcoxon signed-rank tests comparing control and ITI sessions P = 0.02, control and peri-stimulus sessions P = 0.009, ITI and peri-stimulus sessions P = 0.39.

Our study reveals the key role of the ACC in detecting prediction errors and in rapidly shaping subsequent behaviour. Critically, the mismatch responses we found did not resemble classical negative reward signals such as those found in dopamine neurons. Instead, this signal resembles an outcome prediction error[31] which relates to errors in predicting outcomes based on the currently believed rules of the task. The origins of the prediction signal is currently unclear, and may reach the ACC from other prefrontal areas[13,20,42]. However, the ACC itself robustly represents the ongoing task rules as well as external stimuli, and thus may compute the prediction error autonomously.

There is growing evidence to support the role of local inhibitory circuits in the comparison of predicted and observed events[36,43,44]. However, given the opposite effect of VIP activation observed on visuo-motor prediction errors[36] and our results, there may not exist a canonical cortical prediction error circuit, but rather, a variety of circuits supporting various kinds of comparisons[45,46].

Where is the prediction error signal sent to in order to enable the behavioural changes? We have found that the striatum is an unlikely candidate[47], despite being a major projection target of the ACC. Other likely brain regions may include the locus coeruleus, which has been implicated in updating current strategies through norepinephrine signalling[7,8,48], or the dorsal raphe nucleus, which promotes behavioural flexibility through serotonin signalling[49].

Disorders of cognitive flexibility in humans range from excessive transitioning between attentional states in ADHD, to excessive persistence in one or a few repetitive behaviours in ASD[50]. Our results provide a crucial insight regarding the role of the ACC in transitioning between, rather than maintaining, cognitive states. Taken together our results provide causal evidence for a longstanding idea, that the ACC computes prediction errors to guide flexible behaviour.

Author contributions

NC and AGK designed the experiments. NC performed the experiments and analysed the data. MH performed the widefield imaging and contributed to imaging data analysis. DM-J performed the V1 recordings and contributed to imaging and behavioural data analysis. AG built and analysed the RL model. AGK and NC wrote the paper, with contributions from AG. All authors contributed to discussions and commented on the manuscript.

Acknowledgments

We thank John Duncan, Apoorva Bhandari and Sonja Hofer for discussions and comments on the manuscript, and Thomas Mrsic-Flogel, Petr Znamenskiy, Anil Seth, Juan Burrone and members of the Khan laboratory for discussions of the results in this manuscript. We thank Maria Fernandez-Otero for technical assistance and Robert Taylor for assistance in pilot experiments. This work was supported by the Wellcome Trust (AGK, 206222/Z/17/Z), the BBSRC (AGK BB/S015809/1), start-up funds from the CDN, King's College London (AGK), and the MRC CNDD PhD program (MH).

Competing financial interests

The authors declare no competing financial interests

References

1. Monsell, S. Task switching. *Trends in Cognitive Sciences* **7**, 134–140 (2003).
2. Sarafyazd, M. & Jazayeri, M. Hierarchical reasoning by neural circuits in the frontal cortex. *Science* **364**, eaav8911 (2019).
3. Shima, K. & Tanji, J. Role for cingulate motor area cells in voluntary movement selection based on reward. *Science* **282**, 1335–1338 (1998).
4. Heilbronner, S. R. & Hayden, B. Y. Dorsal Anterior Cingulate Cortex: A Bottom-Up View. *Annu Rev Neurosci* **39**, 149–170 (2016).
5. Shenhav, A., Cohen, J. D. & Botvinick, M. M. Dorsal anterior cingulate cortex and the value of control. *Nat Neurosci* **19**, 1286–1291 (2016).
6. Kolling, N. *et al.* Value, search, persistence and model updating in anterior cingulate cortex. *Nat Neurosci* **19**, 1280–1285 (2016).
7. Tervo, D. G. R. *et al.* Behavioral variability through stochastic choice and its gating by anterior cingulate cortex. *Cell* **159**, 21–32 (2014).
8. Tervo, D. G. R. *et al.* The anterior cingulate cortex directs exploration of alternative strategies. *Neuron* **109**, 1876-1887.e6 (2021).
9. Williams, Z. M., Bush, G., Rauch, S. L., Cosgrove, G. R. & Eskandar, E. N. Human anterior cingulate neurons and the integration of monetary reward with motor responses. *Nat Neurosci* **7**, 1370–1375 (2004).
10. O'Reilly, J. X. *et al.* Dissociable effects of surprise and model update in parietal and anterior cingulate cortex. *Proceedings of the National Academy of Sciences* **110**, E3660–E3669 (2013).
11. Narayanan, N. S., Cavanagh, J. F., Frank, M. J. & Laubach, M. Common medial frontal mechanisms of adaptive control in humans and rodents. *Nat Neurosci* **16**, 1888–1895 (2013).
12. Kennerley, S. W., Behrens, T. E. J. & Wallis, J. D. Double dissociation of value computations in orbitofrontal and anterior cingulate neurons. *Nat Neurosci* **14**, 1581–1589 (2011).
13. Wilson, R. C., Takahashi, Y. K., Schoenbaum, G. & Niv, Y. Orbitofrontal Cortex as a Cognitive Map of Task Space. *Neuron* **81**, 267–279 (2014).
14. Vertechi, P. *et al.* Inference-Based Decisions in a Hidden State Foraging Task: Differential Contributions of Prefrontal Cortical Areas. *Neuron* (2020) doi:10.1016/j.neuron.2020.01.017.
15. Schuck, N. W., Cai, M. B., Wilson, R. C. & Niv, Y. Human Orbitofrontal Cortex Represents a Cognitive Map of State Space. *Neuron* **91**, 1402–1412 (2016).
16. Hayden, B. Y., Pearson, J. M. & Platt, M. L. Neuronal basis of sequential foraging decisions in a patchy environment. *Nat Neurosci* **14**, 933–939 (2011).
17. Kolling, N., Behrens, T. E. J., Mars, R. B. & Rushworth, M. F. S. Neural mechanisms of foraging. *Science* **336**, 95–98 (2012).
18. Ito, S., Stuphorn, V., Brown, J. W. & Schall, J. D. Performance monitoring by the anterior cingulate cortex during saccade countermanding. *Science* **302**, 120–122 (2003).
19. Quilodran, R., Rothé, M. & Procyk, E. Behavioral Shifts and Action Valuation in the Anterior Cingulate Cortex. *Neuron* **57**, 314–325 (2008).
20. Bartolo, R. & Averbeck, B. B. Prefrontal Cortex Predicts State Switches during Reversal Learning. *Neuron* **106**, 1044-1054.e4 (2020).
21. Banerjee, A. *et al.* Value-guided remapping of sensory cortex by lateral orbitofrontal cortex. *Nature* **585**, 245–250 (2020).
22. Klein-Flügge, M. C., Bongioanni, A. & Rushworth, M. F. S. Medial and orbital frontal cortex in decision-making and flexible behavior. *Neuron* (2022) doi:10.1016/j.neuron.2022.05.022.
23. Achterberg, J. *et al.* A One-Shot Shift from Explore to Exploit in Monkey Prefrontal Cortex. *J. Neurosci.* **42**, 276–287 (2022).
24. Harlow, H. F. The formation of learning sets. *Psychol Rev* **56**, 51–65 (1949).
25. Akam, T. *et al.* The Anterior Cingulate Cortex Predicts Future States to Mediate Model-Based Action Selection. *Neuron* **109**, 149-163.e7 (2021).

26. Rushworth, M. F. S. & Behrens, T. E. J. Choice, uncertainty and value in prefrontal and cingulate cortex. *Nat Neurosci* **11**, 389–397 (2008).
27. Warren, C. M., Hyman, J. M., Seamans, J. K. & Holroyd, C. B. Feedback-related negativity observed in rodent anterior cingulate cortex. *J Physiol Paris* **109**, 87–94 (2015).
28. Zhang, S. *et al.* Organization of long-range inputs and outputs of frontal cortex for top-down control. *Nat. Neurosci.* **19**, 1733–1742 (2016).
29. Keller, G. B. & Mrsic-Flogel, T. D. Predictive Processing: A Canonical Cortical Computation. *Neuron* **100**, 424–435 (2018).
30. Bastos, A. M. *et al.* Canonical microcircuits for predictive coding. *Neuron* **76**, 695–711 (2012).
31. Alexander, W. H. & Brown, J. W. Medial prefrontal cortex as an action-outcome predictor. *Nat Neurosci* **14**, 1338–1344 (2011).
32. Kuchibhotla, K. V. *et al.* Parallel processing by cortical inhibition enables context-dependent behavior. *Nat. Neurosci.* **20**, 62–71 (2017).
33. Pi, H.-J. *et al.* Cortical interneurons that specialize in disinhibitory control. *Nature* **503**, 521–524 (2013).
34. Pfeffer, C. K., Xue, M., He, M., Huang, Z. J. & Scanziani, M. Inhibition of inhibition in visual cortex: the logic of connections between molecularly distinct interneurons. *Nat. Neurosci.* **16**, 1068–1076 (2013).
35. Lee, S., Kruglikov, I., Huang, Z. J., Fishell, G. & Rudy, B. A disinhibitory circuit mediates motor integration in the somatosensory cortex. *Nat. Neurosci.* **16**, 1662–1670 (2013).
36. Attinger, A., Wang, B. & Keller, G. B. Visuomotor Coupling Shapes the Functional Development of Mouse Visual Cortex. *Cell* **169**, 1291-1302.e14 (2017).
37. Bissonette, G. B., Powell, E. M. & Roesch, M. R. Neural structures underlying set-shifting: roles of medial prefrontal cortex and anterior cingulate cortex. *Behav Brain Res* **250**, 91–101 (2013).
38. Botvinick, M. M. Conflict monitoring and decision making: reconciling two perspectives on anterior cingulate function. *Cogn Affect Behav Neurosci* **7**, 356–366 (2007).
39. Passingham, R. E. *et al.* Attention to action. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences* **351**, 1473–1479 (1996).
40. Procyk, E., Tanaka, Y. L. & Joseph, J. P. Anterior cingulate activity during routine and non-routine sequential behaviors in macaques. *Nat Neurosci* **3**, 502–508 (2000).
41. Hadland, K. A., Rushworth, M. F. S., Gaffan, D. & Passingham, R. E. The anterior cingulate and reward-guided selection of actions. *J Neurophysiol* **89**, 1161–1164 (2003).
42. Durstewitz, D., Vittoz, N. M., Floresco, S. B. & Seamans, J. K. Abrupt Transitions between Prefrontal Neural Ensemble States Accompany Behavioral Transitions during Rule Learning. *Neuron* **66**, 438–448 (2010).
43. Jordan, R. & Keller, G. B. Opposing Influence of Top-down and Bottom-up Input on Excitatory Layer 2/3 Neurons in Mouse Primary Visual Cortex. *Neuron* **108**, 1194-1206.e5 (2020).
44. Eliades, S. J. & Wang, X. Neural substrates of vocalization feedback monitoring in primate auditory cortex. *Nature* **453**, 1102–1106 (2008).
45. Hertäg, L. & Clopath, C. Prediction-error neurons in circuits with multiple neuron types: Formation, refinement, and functional implications. *Proceedings of the National Academy of Sciences* **119**, e2115699119 (2022).
46. Hertäg, L. & Sprekeler, H. Learning prediction error neurons in a canonical interneuron circuit. *Elife* **9**, e57541 (2020).
47. Spellman, T., Svei, M., Kaminsky, J., Manzano-Nieves, G. & Liston, C. Prefrontal deep projection neurons enable cognitive flexibility via persistent feedback monitoring. *Cell* **184**, 2750-2766.e17 (2021).
48. Aston-Jones, G. & Cohen, J. D. An integrative theory of locus coeruleus-norepinephrine function: adaptive gain and optimal performance. *Annu Rev Neurosci* **28**, 403–450 (2005).
49. Clarke, H. F., Dalley, J. W., Crofts, H. S., Robbins, T. W. & Roberts, A. C. Cognitive inflexibility after prefrontal serotonin depletion. *Science* **304**, 878–880 (2004).

50. Uddin, L. Q. Cognitive and behavioural flexibility: neural mechanisms and clinical considerations. *Nature Reviews Neuroscience* 1–13 (2021) doi:10.1038/s41583-021-00428-w.

METHODS

All experimental procedures were carried out in accordance with the institutional animal welfare guidelines and licensed by the UK Home Office.

Surgical Procedures

For all surgeries, mice were anaesthetised using isoflurane, at 4% concentration for induction and at 0.5-1% for maintenance. Additional drugs were used to provide analgesia (Metacam 5mg/kg), anti-inflammatory effects (dexamethasone 3.8mg/kg), and to reduce mucus secretions (Atropine 0.08mg/kg). Eye-cream (Maxitrol) was applied to the eyes to prevent drying and body temperature was maintained at 37°C using a heating mat and rectal temperature probe (Harvard Apparatus). Injections of antibiotic (Betamox 120mg/kg) and analgesia (methadone hydrochloride 10mg/kg) were given before the withdrawal of anaesthesia, and further analgesia was given daily for 1-2 days during recovery of the animal.

For the imaging experiments, 10 C57BL/6-VIP$^{tm1(cre)zjh}$ mice (The Jackson Laboratory, strain # 010908, P42-49, 6 males and 4 females) were used. A circular piece of scalp was removed and the underlying skull was cleaned. Small holes were drilled in the skull above injection sites, located using stereotaxic co-ordinates. Injections of a mixture of viruses expressing GCaMP7f (pGP-AAV9-syn-jGCaMP7f-WPRE, Addgene) and ChrimsonR (pAAV5-syn-FLEX-rc[ChrimsonR-tdTomato], Addgene) were made in the anterior cingulate cortex (ACC, +0.9mm AP, -1.3 rising to -0.8mm DV) of either the left or right hemisphere (+/- 0.55mm ML), using glass pipettes and a pressure micro-injection system (Picospritzer III, Parker). An injection of Cholera Toxin Subunit B (recombinant) conjugated to Alexa Fluor 647 (ThermoFisher) was made in the striatum of the same hemisphere (+1.2mm AP, +/- 1.5mm ML, -2.5 rising to -2.0mm DV) to retrogradely label cells projecting to the striatum. 8 out of the 10 mice had good enough quality z-stacks in the ACC to visualise CTB-Alexa647 and were used to identify striatal projecting ACC neurons.

A circular craniotomy (diameter = 3mm) was made above the ACC imaging site, with a centre 300µm posterior to ACC injections. A 1.5mm x 1.5mm x 1.5mm right-angled microprism with a reflective hypotenuse (Tower Optical) fixed to a glass coverslip using ultraviolet light-cured glue (Thorlabs) was slowly lowered into the craniotomy, with the vertical face closest to the injection site. The glass coverslip was fixed in place using cyano-acrylic glue (Loctite) and a custom machined aluminium head-plate was cemented onto the skull using dental cement (C&B Superbond). Imaging and behavioural training started approximately three weeks after surgery.

For the widefield imaging experiments, four male wildtype C57BL/6;129-*Nrxn1*^{tm1Sud} mice from the ages of 21 to 27 weeks were used (all mice were wildtype C57BL/6). At least four weeks before surgery mice were given an intravenous injection of AAV-PHP.eB-hSyn1-GCaMP7f (Zurich vector core). For the surgery, the skin overlying the skull was removed and the edges of the skin were secured with tissue adhesive (Vetbond, 3M). The overlying connective tissue on the skull was removed and a layer of transparent dental cement (C&B Superbond) was applied to cover all exposed skull and to secure a custom aluminium headplate. Following this 5 layers of 2.5ul of cyanoacrylate glue (Zap-A-Gap CA+, Pacer Technology) were thinly applied onto the cement to increase the transparency. After recovery from surgery for at least 5 days mice began habituation and behavioural training.

For the optogenetic silencing experiments, 8 transgenic mice (P42-49, 4 males and 4 females) expressing Channelrhodopsin-2 in parvalbumin-expressing interneurons were used, obtained by crossing Gt(ROSA)26Sor$^{tm32(CAG-COP4*H134R/EYFP)Hze}$ and Pvalb$^{tm1(cre)Arbr}$ mice (The Jackson Laboratory, strain # 012569 and # 017320 respectively). Two small holes were drilled above the ACC and prelimbic cortex

(PL) of each hemisphere. Dual-core cannulae with bilateral optical fibres (Thorlabs), each with a diameter of 200µm and 0.39NA, cut to a length of < 3mm, were implanted in the ACC (+0.9mm AP, -1.2mm DV, +/-0.35mm ML) and PL (+2.6mm AP,-1.25mm DV, +/- 0.35mm ML), and the stainless steel ferrules were bonded to the skull using dental cement (C&B Superbond), along with a custom machined head-plate. PL implants were inserted at a 25° angle (relative to vertical) through holes drilled 0.8mm anterior to PL. For light-only control experiments, 5 Pvalb$^{tm1(cre)Arbr}$ mice (3 males and 2 females) were implanted with cannulae above ACC only. After recovery from surgery for at least 5 days mice began habituation and behavioural training.

Imaging

Two-photon imaging was performed using a custom-built resonant scanning two-photon microscope (Cosys) and a Chameleon Vision S laser (Coherent) at 930nm using a 16X, .8NA objective (Nikon). Images were acquired using a 12 KHz resonant scanner (Cambridge Technology) and an FPGA module (PXIe-7965R FlexRIO, National Instruments). Two-photon calcium imaging of GCaMP7f-labelled neurons in the ACC was performed across 40 training sessions and 48 full task-switching sessions in these 10 mice. The microprism depth, injection coordinates and cell morphology confirmed that the imaging sites were located in layer 5. Multi-plane imaging was performed using a piezoelectric objective scanner (Physik Instrumente). Depending on the depth of GCaMP7f expression, each imaging volume consisted of either 6 or 8 imaging planes, 40µm apart, giving an effective imaging rate of 6.4 or 4.8Hz per volume respectively.

Mice were trained first in the visual discrimination task, then had at least 3 training sessions in the visual-odour block switching task. Once mice had learned the switching task, at least 3 recordings of the mice performing the task were taken per mouse. Before each recording session the same imaging site was found by matching anatomical landmarks.

After all in-vivo imaging data had been collected, a final high-quality image stack was acquired under anaesthesia. Subcutaneous injections of ketamine (100mg/kg) and xylazine (16mg/kg) were used to induce anaesthesia, with further injections of ketamine to maintain anaesthesia if necessary. Eye-cream (Maxitrol) was used to prevent drying, and body temperature was maintained using a heating pad.

Widefield imaging was performed on a custom built inverted tandem lens macroscope (Cosys), with two photographic lenses (AF-S NIKKOR 85mm f/1.8G lens and AF NIKKOR 50mm f/1.4D Lens). The brain was illuminated with interleaved collimated blue (470nm, Thorlabs M470L4) and violet light (405nm, Thorlabs M405L4) at an irradiance of ~0.03mW/mm^2. Images were recorded with a CMOS camera (Point Grey Research Grasshopper3) at frame rate of 54Hz. LEDs and camera frame acquisition were triggered using a digital microprocessor (Teensy 3.2).

The widefield video underwent motion correction and the brain images were aligned within and across mice by manual rigid alignment to a number of anatomical landmarks. The ΔF/F was computed for each pixel by taking the difference between F and F_0, and dividing by F_0, where F_0 was the mean value across the entire session. Traces were filtered with a 0.0033 Hz high-pass second order Butterworth filter, and an additional 7Hz lowpass filter was applied to the violet trace. To correct for haemodynamic artefacts, a scaled version of the violet illumination trace was subtracted from the blue illumination trace for each pixel. This scaling factor was found by regressing the violet trace onto the blue trace. To account for overt movement-related brain activity, we fit a ridge regression model to the data, predicting brain activity from running speed and licking. We then subtracted this predicted activity and all subsequent analysis was performed on the model residuals.

Behavioural training

The behaviour apparatus and training was similar to previous studies[51,52]. Mice were trained on a visual discrimination task for up to two weeks, until discrimination performance reached threshold, before training them on the switching task (see below). Mice had free access to water but were food restricted to maintain at least 85% of their free-feeding body weight (typically 85-90%, 2-3g of standard food pellets per animal per day). A reward delivery spout was positioned near the snout of the mouse, and licks to this spout were detected using a piezo disc sensor and custom electronics. The reward was a 10% solution of soy milk powder (SMA Wysoy) delivered by opening a pinch valve (NResearch) controlled through custom electronics. The mouse's running speed on the cylinder was measured using an incremental rotary encoder (Kübler). Two luminance-corrected monitors (luminance meter LS-100, Konica Minolta) positioned at 45° angles and 25cm distance to the mouse delivered visual stimuli.

Animals were habituated to handling and gentle restraint over two to three days, before they were head-fixed and trained to run on a polystyrene cylinder (20cm diameter) for one to four days. This period was also used to find suitable imaging sites. After the habituation phase, mice performed one behaviour session in which the movement of the gratings was linked to the mouse's movement on the wheel. Subsequently, mice were trained to self-initiate trials by sustained running on the wheel for at least 2.8s and an added random duration drawn from an exponential distribution with mean 0.4s. At this point one of two drifting sinusoidal visual gratings were randomly presented, drifting in the opposite direction to the direction of running, with a fixed spatial and temporal frequency of 0.1 cycles per degree and 2Hz respectively. The rewarded and unrewarded gratings were oriented +/-20° relative to vertical, symmetrically on both screens. When the rewarded grating was displayed the mouse could trigger the delivery of a reward, a drop of soya milk, by licking the spout during the 'reward period', lasting from 1.5s after the appearance of the grating to its disappearance, maximum 1.53 s into the reward period. This was recorded as a 'hit'. If the mouse did not lick during this period, the trial was recorded as a 'miss', and a drop of soy milk was delivered shortly before the disappearance of the grating. When the unrewarded grating was presented, a single lick or more at any time until the stimulus disappearance was recorded as a 'false alarm', triggering a time-out period of 4s in which the unrewarded grating remained on screen, and any further licks restarted the time-out. During early training the probability of unrewarded trials was occasionally increased transiently up to 0.8 to discourage erroneous licking. All mice learned the visual discrimination task in 5-10 days, with post-learning defined as three consecutive days of discrimination with a behavioural d-prime score of 2 or above. Behavioural d-prime was calculated as: $bd' = \Phi^{-1}(H) - \Phi^{-1}(F)$, where Φ^{-1} is the normal inverse cumulative distribution function, H is the rate of hit trials, and F is the rate of false alarm trials.

Once mice had learned the visual discrimination task, they were trained in odour discrimination. After the same randomised period of sustained running, one of two odour stimuli were presented to the mouse via a polyethylene tubing positioned above the snout of the mouse. Odours were delivered through a custom built flow dilution olfactometer calibrated with a mini PID (Aurora) at 10-20% saturated vapour concentration of two solutions, 10% soy milk (rewarded odour) and 10% soy milk with 0.1% p-Cymene mixture (unrewarded odour). The odour task structure was identical to the visual task.

Once animals were discriminating the odours accurately (typically after 30-40 trials), they were trained to switch between blocks of the olfactory and visual discrimination task. Mice typically learned to switch successfully in 1-3 days. In the olfactory blocks, 70% of odour stimuli were preceded with one of the two same visual gratings featured in the visual discrimination task (fixed duration of 1.8s, with an identical onset delay distribution as in the visual block). In this instance neither grating was

rewarded or punished, and mice learned to ignore these irrelevant gratings while accurately discriminating the odours, which were presented after the irrelevant visual grating (delay between visual grating offset and odour onset 1.5s, plus an added random duration drawn from an exponential distribution with mean 0.2s in imaging sessions). In initial switching training sessions, a reward was delivered at the end of a rewarded grating in a visual block if the mouse had failed to lick, giving a clear indication that the grating was now relevant. By the end of early training, and for all data in this paper except Extended Data Fig.4e, this feature was removed, requiring mice to switch between blocks through noticing unexpected stimuli or outcomes alone. Block switches occurred automatically when a mouse had demonstrated a > 80% discrimination performance to the relevant stimuli (visual gratings in visual block, odours in odour block) over the last 30 trials of a block. Additionally in odour blocks mice were required to have successfully ignored all irrelevant visual gratings over the previous 10 trials. Blocks typically contained 30 to 40 trials. Mice were deemed to have learned the switching task when they could complete sessions at these parameters with at least 3 repeats of each block type.

In order to compare the speed of switching between blocks, the selection of visual stimuli immediately after a block transition was made uniform during a transition period. In the first trials of a visual block only the rewarded grating (visual stim 1) was shown, while in the first trials of an odour block the otherwise rewarded grating (visual stim 1) preceded every odour stimulus as the irrelevant grating. Odour stimuli selection itself was kept random. When a mouse responded correctly to these grating stimuli on three consecutive trials, by licking in the visual block and by withholding licks in the odour block, this transition period ended and the block continued with the normal 50% probabilities of visual grating identities. These transition periods were used in all behaviour sessions except muscimol silencing experiments (Extended Data Fig. 1a) and light-only controls (see below), in which case either visual stimulus was presented from the start of the block with 50% probability.

Reinforcement learning model

We modelled the experimental protocol of stimuli and rewards based on mouse actions as a reinforcement learning environment. The environment was written in Python with the OpenAI Gym interface for ease of use with other agents. The environment had 5 states, two for visual cues, two for olfactory cues and an end of trial cue, and two possible actions that the agent could perform, lick and no-lick. Each trial comprised a number of steps.

In a visual block, each trial had 2 steps: in step 1, a needless lick (to the previous end of trial) was punished (-1 assuming an internal cost), and one of 2 visual stimuli were shown; in step 2: a lick led to reward (+1 corresponding to soy milk drop) if visual cue 1 was presented, or punishment (-1 corresponding to experimental timeout) if visual cue 2 was presented, and the end of trial was indicated. In an olfactory block, a trial had 2 time steps in 30% of the trials, corresponding to trials without irrelevant visual stimuli, and 3 time steps in 70% of the trials, corresponding to trials with irrelevant visual stimuli. In the 2 time steps case, in step 1, a needless lick was punished (-1), and one of 2 odour stimuli were given. In step 2, a lick led to reward (+1) if odour 1 was delivered, or punishment (-1) if odour 2 was delivered in step 1, and end of trial was indicated. In the 3 time step case, in step 1, a needless lick was punished (-1), and one of 2 visual stimuli were shown. In step 2, a needless lick was punished (-1), and one of 2 odour stimuli were delivered. In step 3, a lick led to reward (+1) if odour 1 was delivered, or punishment (-1) if odour 2 was delivered in step 2, and end of trial was indicated. No lick always led to 0 reward.

As in the experiment, after every transition, the trial type was fixed such that the rewarded visual stimulus was always shown in step 1 of either block until the agent performed 3 correct responses i.e.

licking in transition to visual block, no licking in transition to olfactory block. Block transitions occurred after the agent achieved at least 80% correct responses in past 30 trials, and 100% correct responses in ignoring past 10 irrelevant visual stimuli, with a minimum of 30 trials after initial shaping trials after a block transition. A lick for cue 1 or a no lick for cue 2 in the final reward step counted as correct response, while requiring no lick in all other time steps.

Basic RL agent (Tabular SARSA algorithm): Each of the 2 visual and 2 olfactory cues, as well as the end-of-trial cue, was considered a state. The possible actions were lick and no-lick. A Q-value $Q(s,a)$ table was constructed with entries for each combination of 4 states and 2 actions, denoted by s and a, initialized to zeros. The Q-value $Q(s,a)$ represented the expected total reward till the end of the trial, given the cue s and taking action a at the current step. The Q-value for the end of trial cue was set to 0. As each cue s was encountered, the agent responded with an action a according to an ϵ-greedy policy. i.e. a random action was taken with probability ϵ, otherwise the action that yielded the maximum Q-value for the current cue s was taken. The Q-table was updated as per the SARSA (State-Action-Reward-State-Action) algorithm using the temporal difference (TD) error

$$\delta \equiv r + Q(s',a') - Q(s,a),$$

multiplied by a learning rate α

$$Q(s,a) \leftarrow Q(s,a) + \alpha\delta$$

Cues s', s and actions a', a represent the current, previous time steps respectively.

Belief State RL agent: Two Q-value tables were constructed and initialized to zero, corresponding to the visual and olfactory blocks, each of size 4 states by 2 actions. The agent also had a belief b about being in a visual block v versus an olfactory block o, which was represented as a discrete probability distribution $b \equiv (p(v), p(o))$. At every step of the trial, the agent assumed that the current block was either visual or olfactory depending on which probability $p(v)$ or $p(o)$ was higher, and took an action according to an ϵ-greedy policy based on the Q-table corresponding to the assumed block. This Q-table was updated similar to the basic agent using the TD-error δ multiplied by a learning rate α.

At the end of every trial, a block mismatch signal $\chi \equiv d - b$ was computed as the difference between the detected block d (represented as (1,0) or (0,1) for visual or olfactory block respectively depending on whether the cue just before the end of the trial was a visual or odor cue), and the agent's belief b. A noise-corrupted version of this block mismatch signal $\chi' = \chi(1 + \beta\xi)$ was computed, where ξ was a Gaussian random variable with mean 0 and variance 1, and β was a noise factor parameter. The agent's belief was updated as:

$$b' = b + \zeta\chi'$$

where χ' was multiplied by a belief switching rate ζ, and then each component of b' was clipped to be greater than 0 and clipped b' was normalized to yield a probability distribution b which served as the belief of the next trial.

Model fitting: Each model was simulated with learning via SARSA for N time steps. We used the simulated behaviour data from $N/2$ time steps onwards, by which time the agent had learnt the task. For fitting, we minimized the root mean squared error (RMSE) between the experimental and simulated $p(lick|cue)$. For all fits, we performed a global grid search within reasonable parameter ranges, followed by a local minimization starting from the best parameter sets obtained from the grid

search. The fits were performed using 5-fold cross validation, where for each fold, we fit the parameters of the model to 4/5th of the data and tested on the held out 1/5th of the data. To select between the two models, RMSE mean $\pm$ SD on training and test splits were calculated across 5 seeds. These were 0.145960 $\pm$ 0.004752 and 0.182556 $\pm$ 0.014442 for basic RL, and 0.082383 $\pm$ 0.00618 and 0.138593 $\pm$0.015528for belief-state RL.

For the basic RL model, we fit 2 parameters: exploration rate ϵ and the learning rate α. This model was unable to reproduce the rapid block transitions observed in the data, since both the exploration rate and the learning rate needed to be very high to rapidly explore actions and learn a different reward structure after a transition, but a high exploration rate was inconsistent with the steady-state experimental data. For the belief-state RL model, we fit 4 parameters: belief switching rate ζ, the noise factor β for the mismatch signal, exploration rate ϵ, and a different exploration term ϵ' to account for enhanced licking to visual cue 2 in the olfactory block. The learning rate α was fixed at 0.1 and was not fit because the belief switching rate played a much stronger role in rapid switching between blocks.

Simulated $p(lick|cue)$ using the best fit parameters (on the full dataset) for each model for one seed are shown in Figure 1g middle and bottom. The mismatch signal for one-shot versus slower switches made by the agent, shown in Fig. 5c, was computed as the sum of the absolute values of the two components of the mismatch signal χ at the end of the first trial following the block transition. For each of both types of transitions, only 70 transitions were chosen randomly from the simulated data, similar to the number of transitions in the experimental data, for plotting and significance testing.

For fitting the behaviour during ACC silencing, which was a separate experimental dataset, first we fitted the default 4 parameters for the belief-state RL model on the behaviour data without ACC silencing. Then keeping these parameters fixed, we fitted the behaviour data with ACC silenced, using two parameters – factors which were multiplied by the mismatch signal for odour and visual trials. These factors signified how much the mismatch signal reduced on silencing the ACC. Since these fits were not for model selection, all of the data was fit, minimizing mean RMSE across 5 seeds.

Fitted parameters are shown in Table 1.

Type of model / experimental data	Fixed parameters	Fitted parameters	Mean $\pm$ SD of RMSEs across 5 seeds
Basic RL on primary dataset	*N* = 2,000,000 (1,000,000 during fitting)	ϵ, α = 0.25, 0.9	0.143954 $\pm$ 0.001384
Belief-state RL on primary dataset	*N* = 500,000 α = 0.1	$\zeta, \beta, \epsilon, \epsilon'$ = 0.695, 2.06 , 0.075, 0.325	0.075087 $\pm$ 0.000368
Belief-state RL on ACC silencing dataset with ACC not silenced	N=500,000 α = 0.1	$\zeta, \beta, \epsilon, \epsilon'$ = 0.713, 2.05, 0.083, 0.474	0.053966 $\pm$ 0.000826

Belief-state RL on ACC silencing dataset with ACC silenced	N=500,000 $\alpha = 0.1$ $\zeta, \beta, \epsilon, \epsilon' =$ 0.713, 2.05, 0.083, 0.474	ACC inhibition factors on mismatch for visual and odour trials = 0.220, 0.948	0.074892 ± 0.001570

Table 1: Best-fit parameters for each model.

Imaging data analysis

Image stacks were corrected for motion, and regions of interest (ROIs) were selected for each active cell in each session using Suite2p[53]. Each site yielded between 129 and 925 cells, median = 499 cells. Raw fluorescence time series F(t) were obtained for each cell by averaging across pixels within each ROI. Baseline fluorescence $F_0(t)$ was computed by smoothing F(t) (causal moving average of 0.75s) and taking the 2.5th percentile of the smoothed data. The change in fluorescence relative to baseline, ΔF/F, was computed by taking the difference between F and F_0, and dividing by F_0.

To analyse responses to the odour prediction mismatch trials, neuronal activity was aligned to the end of the preceding visual grating plus 2.0s, the average delay to the odour stimuli onset. Mismatch neurons were identified by comparing activity between three conditions in the window 0 to 1.5s aligned to this time: activity in odour prediction mismatch trials (i.e. after the first grating following a switch from an odour block to a visual block, provided the mouse did not already lick to the preceding grating), activity in the same time period following a stable unrewarded visual block trial up to 10 trials before the end of a visual block (when no odour should be expected) and activity during a stable odour block trial (when an odour is expected and received) following a correctly ignored visual grating, using a Wilcoxon rank-sum test with equal numbers of trials for each condition. If multiple trials were collected from each block for the stable visual block and stable odour block conditions, activity was averaged across trials from the same condition and same block. Mismatch neurons were defined as cells with significantly different activity in odour prediction mismatch trials, when compared to both of the other conditions. Positive and negative mismatch neurons were those in which the response to the odour prediction mismatch was largest or smallest of the three conditions respectively. Two other combinations were observed, first with the odour condition significantly higher and no odour condition significantly lower compared to the prediction mismatch condition (99 cells), and second, the reverse of this (163 cells). To test whether activity in mismatch neurons could predict subsequent switching, average activity between 0-1.5s in odour prediction mismatch trials was compared between one-shot switches (in which the mouse responded correctly to the subsequent visual grating) and slower switches (in which the mouse continued to miss at least the next visual grating) using a Wilcoxon signed-rank test.

Decoding analysis (Fig. 2c) was performed by training a binary logistic regression classifier to decode the stimulus or block identity from the vector of ΔF/F values at each frame. Decoding performance was assessed using stratified 5-fold cross validation and taking the mean accuracy across the 5 test sets. Stimuli classes were evenly balanced by randomly subsampling the larger class. We applied an L2 regularisation penalty to reduce overfitting.

For identification of striatal projecting cells, a brief dual channel recording of the imaging planes was acquired before each imaging session at an excitation wavelength of 830nm. Following segmentation, imaged neurons co-expressing Alexa-647 were identified using this recording, and confirmed using a detailed anesthetised dual channel z-stack taken at the end of all imaging sessions. To calculate

confidence intervals for the percentage of mismatch cells in the non-striatal projecting ROIs (Fig. 2i) a percentile bootstrap method was used, resampling with replacement a number of ROIs equivalent to the size of the striatal projecting ROI population 10,000 times. The proportion of mismatch cells was then calculated and the 0.5^{th} and 99.5^{th} percentiles of this distribution of proportions was calculated to obtain the 99% confidence interval.

To assess the proportion of neural activity which was attributable to overt behaviour recorded during our task (Extended Data Fig. 2e-g) a linear model was fit using ridge regression to predict neural activity. The model was constructed by combining multiple sets of variables into a design matrix, to capture signal modulation by the following different task or behavioural events: 2 visual stimuli, 2 odour stimuli, reward delivery, licks, running speed, block type, and an interaction term for visual stimuli and block type. Each stimulus/event variable was structured to capture a time-varying event kernel. Variables therefore consisted of a vector of the relevant stimulus/event, and copies of this vector, each shifted in time by one frame for specific durations. For sensory stimuli, the time-shifted copies ranged up to 2s after the original. For motor events (running and licking) the time-shifted copies spanned the frames from 0.5s before until 2s after the original. The model was fit with 5-fold cross validation and the coefficient of determination (R^2) was calculated based on the predictions of the model on held out data not used during training. We then assessed the predictive power of the behavioural model variables by comparing the R^2 value for the full model to a model without the running and licking predictors.

Optogenetic activation of VIP interneurons during ACC imaging

To perturb VIP cells expressing Chrimson concurrently with 2-photon imaging, a 639nm laser (OBIS, Coherent) was used to deliver light via a 200µm diameter, 0.39 NA optic fibre (Thorlabs) positioned around 3mm from the posterior edge of the microprism, at a 30° angle relative to the coverslip. The laser and stimulus monitors were blanked during the linear phase of the resonant scanner to allow quasi-simultaneous two-photon imaging and optogenetic activation. The effective maximum output power from the optic fibre was 5.3mW. During an optogenetic calibration session in the dark, 5 laser powers (20%, 40%, 60%, 80%, and 100% of maximum) were pseudorandomly applied to the coverslip for 1.5s, with a 5 second interval between each. During switching sessions, two laser epochs were used, with 5.3mW power only. In the peri-stimulus epoch, the laser began 0.1s before the visual stimulus and continued until the end of the visual stimulus. In the inter-trial-interval epoch, the laser began at the offset of each visual stimulus and continued until the onset of the next visual stimulus.

Optogenetic silencing of ACC activity

Once mice had learned the full switching task, optogenetic silencing of ACC neurons was performed by connecting the optic fibre cannulae to a blue LED (470nm, Thorlabs), and delivering light while the mouse performed the task. Before implantation, each optic fibre was confirmed to have an effective power output of >1mW after cutting. Light was delivered either throughout the session (pulsed at 40Hz[54]), 0.5s before each visual stimulus continuing to the end of the stimulus ('peri-stimulus'), or from the end of each visual stimulus until the beginning of the next visual stimulus ('inter-trial-interval', ITI). These three epochs were used to silence both ACC and PL on different days, creating 6 silencing conditions. These conditions were pseudorandomly chosen across consecutive days, with the order different between mice but the time between repeated conditions kept constant, and with control no-light sessions interspersed every third session. For one-trial un-silencing, the light was continuously pulsed throughout the session as above, but this was paused at the end of the last trial in the odour block, and resumed at the start of the second trial of the visual block. In the peri-stimulus and ITI epochs, light power was ramped down for the final 200ms of each pulse. For ACC silencing

during learning, the silencing group included all 8 optogenetic mice, and the non-silenced controls were the 10 imaging mice. Light only controls were performed in wild-type mice not expressing channelrhodopsin (N = 5 mice). No differences were found in these mice in the number of trials to transition between blocks with and without the light stimulation ($P > 0.05$, Wilcoxon rank sum test).

Pharmacological silencing of ACC activity

For the muscimol silencing experiments, 4 male wildtype C57Bl/6j mice were implanted with bilateral infusion cannulae (-1.2mm DV, +0.7mm AP, ± 0.5mm ML, bilaterally). After training in the switching task, mice were infused bilaterally with 300nl muscimol (Sigma, 1µg/ µl) or saline at a rate of 0.25µl/min using a 1µl syringe (Hamilton) and syringe pump (World Precision Instruments SP100IZ) 30 minutes before the start of a switching session. We waited 5 minutes after the syringe pump had finished to allow full infusion of the drug before disconnecting the cannulae.

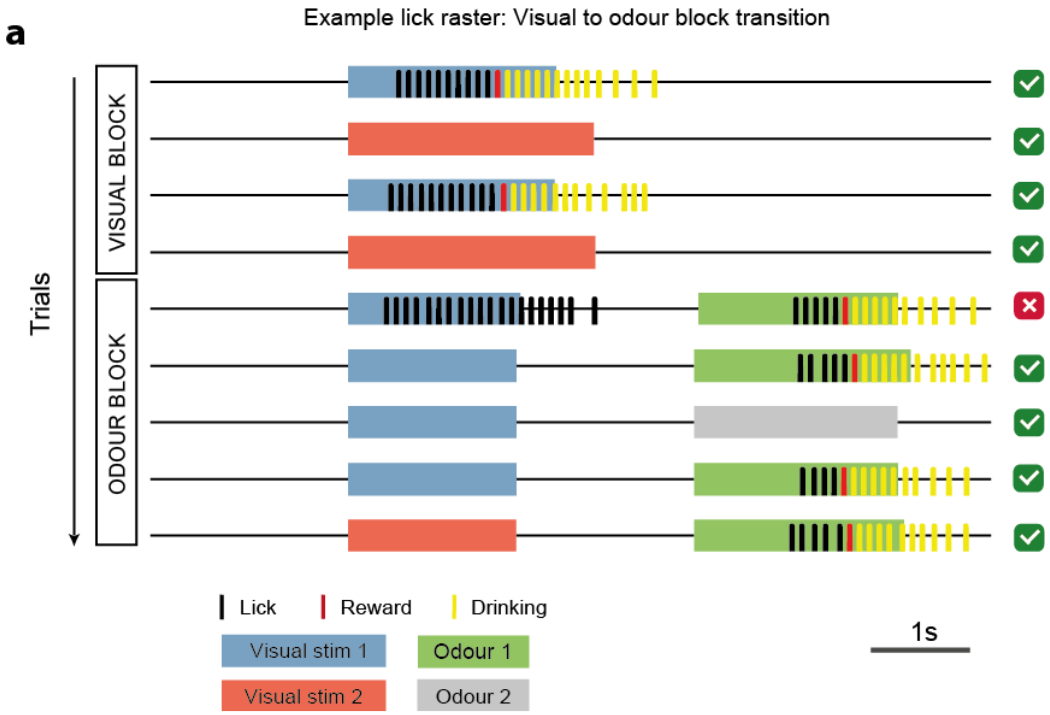
Silencing data analysis

Behavioural d-primes in each silencing condition were calculated by taking performance in stable periods of blocks, outside of transition periods[51,52]. Switching speeds were defined as the number of trials that elapse before a mouse correctly responded to three rewarded visual gratings in a row, either licking to the grating after a switch to a visual block or by ignoring the grating after a switch to an odour block. A 'fluke' switch was defined as a switch in which the mouse correctly licked to the very first rewarded grating in a new visual block, before any evidence of the switch had been received. These were interpreted as exploratory licks, visible in the histograms in Fig. 1e and Fig. 6d at trial 0, and all such transitions were excluded from analysis of switching speeds.

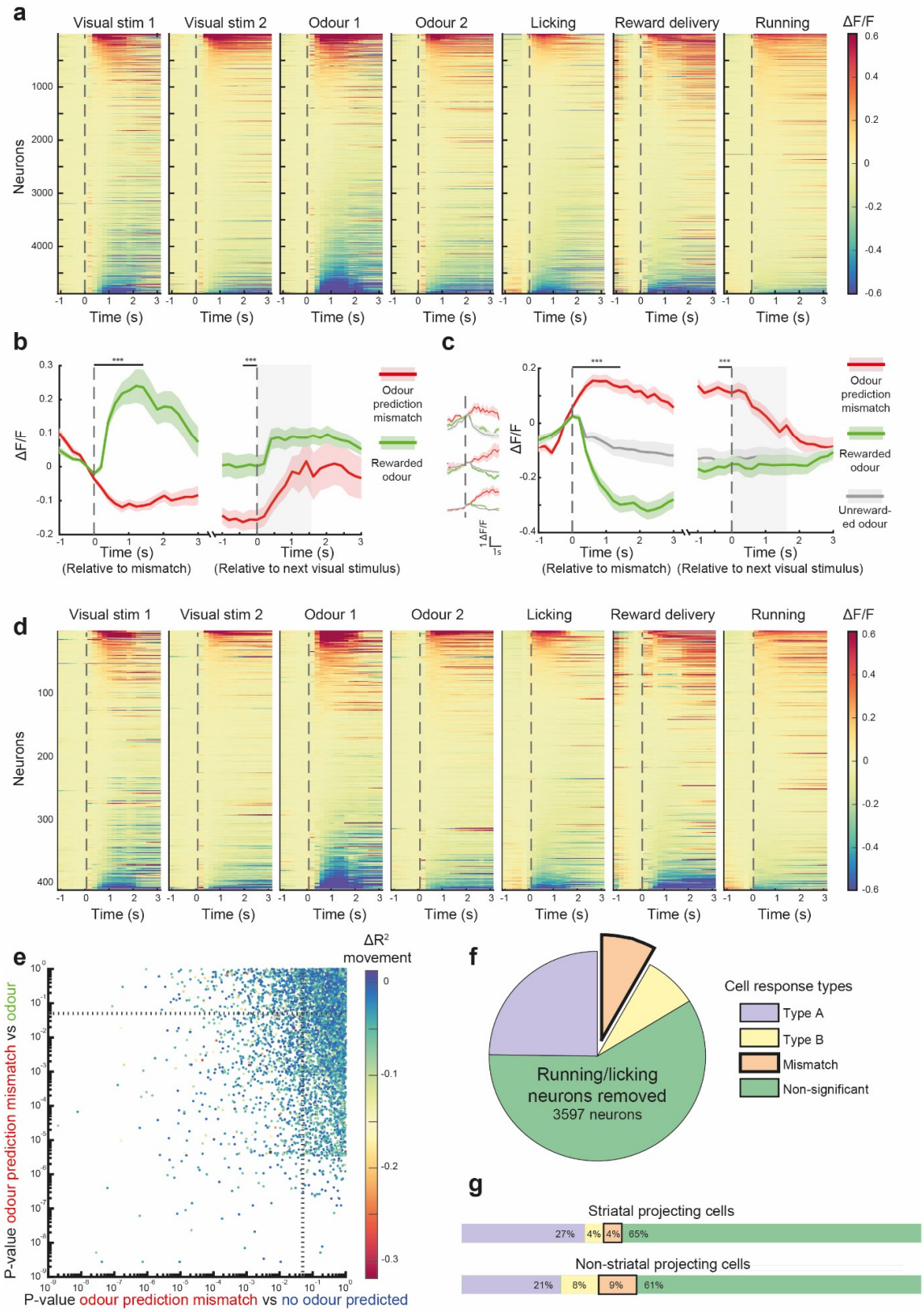
References

51. Poort, J. *et al.* Learning Enhances Sensory and Multiple Non-sensory Representations in Primary Visual Cortex. *Neuron* **86**, 1478–1490 (2015).
52. Poort, J. *et al.* Learning and attention increase visual response selectivity through distinct mechanisms. *Neuron* **110**, 686-697.e6 (2022).
53. Pachitariu, M. *et al.* Suite2p: beyond 10,000 neurons with standard two-photon microscopy. 061507 Preprint at https://doi.org/10.1101/061507 (2017).
54. Li, N. *et al.* Spatiotemporal constraints on optogenetic inactivation in cortical circuits. *Elife* **8**, (2019).

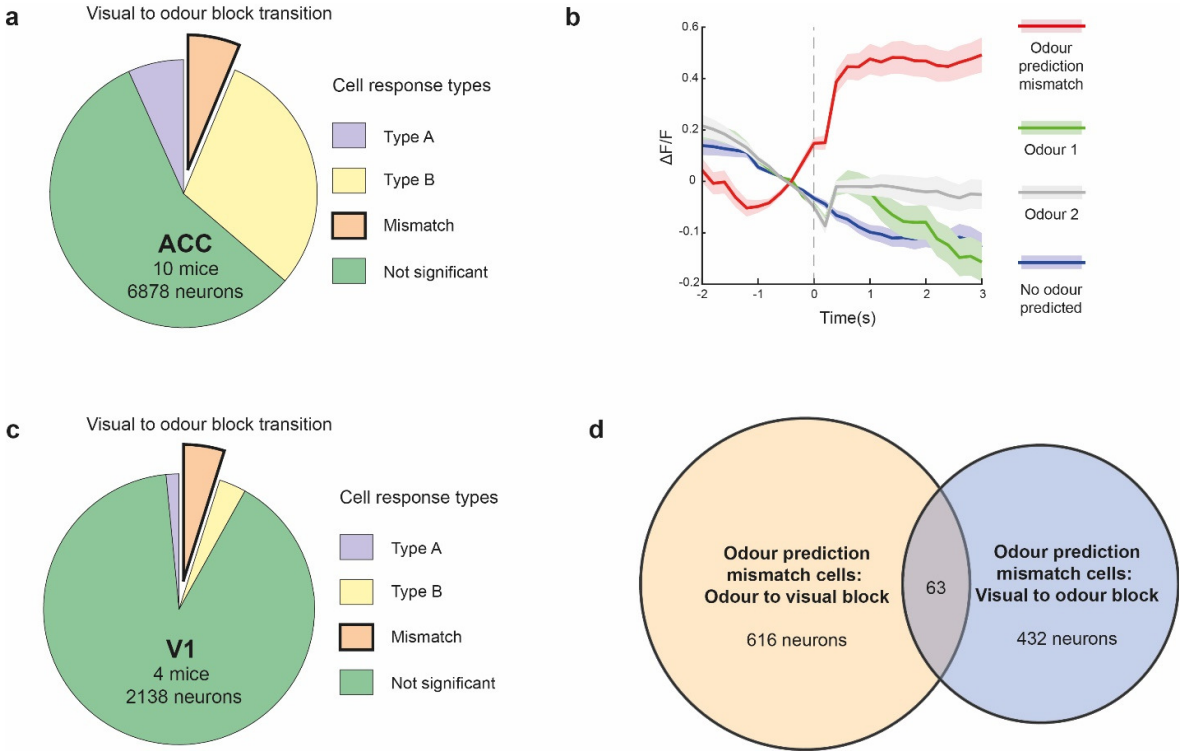
SUPPLEMENTARY FIGURES



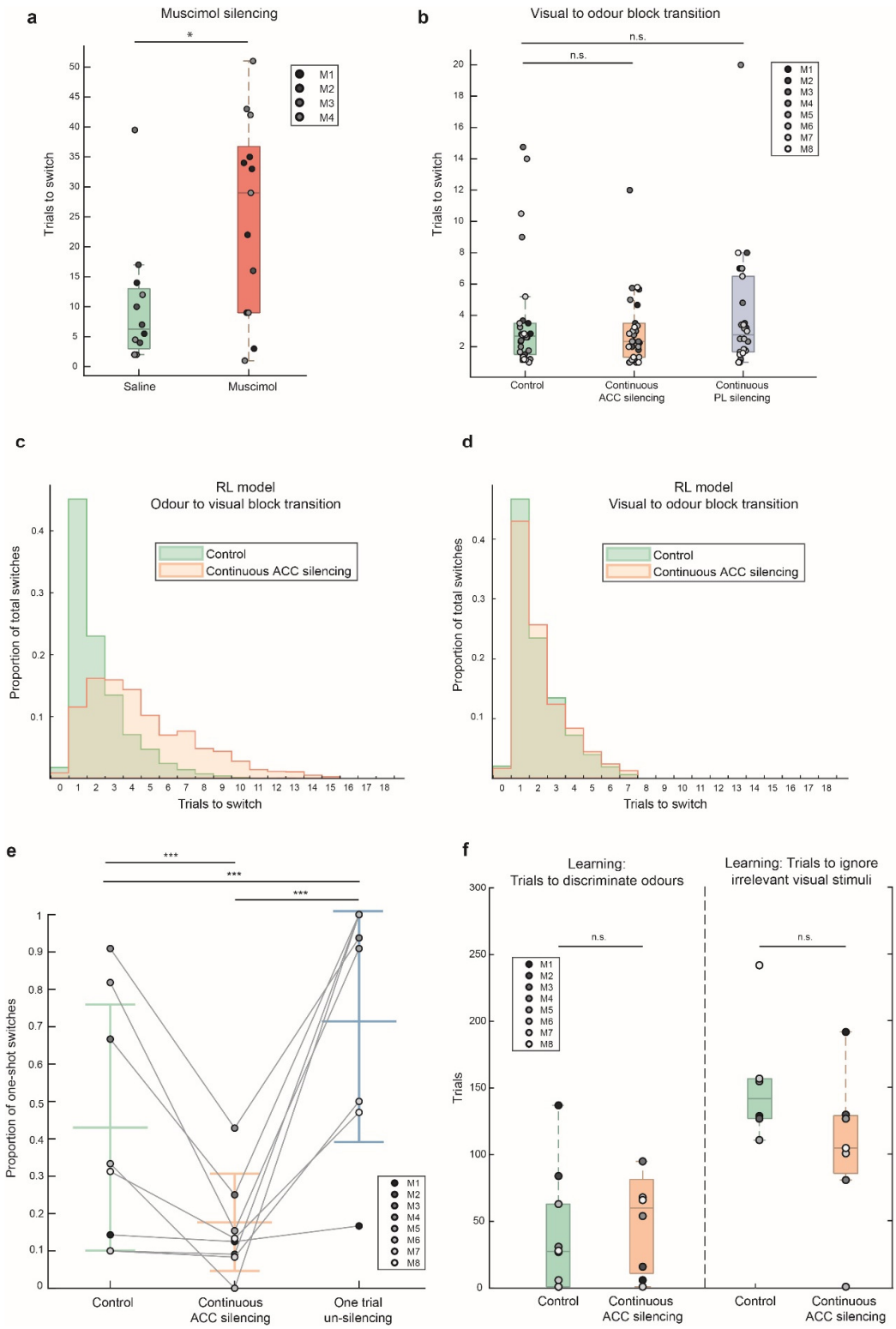
Extended Data Figure 1 | Mice perform rapid task-switching with one-shot block transitions. a. Example behavioural data from a visual to odour block transition (lick raster) showing stimuli, lick and reward times. In the visual block the mouse discriminates both visual gratings, but switches rules after a single error trial and accurately starts ignoring the same visual stimuli while accurately discriminating odour stimuli.


a
Visual stim 1
Visual stim 2
Odour 1
Odour 2
Licking
Reward delivery
Running
ΔF/F
Neurons
Time (s)
b
c
Odour prediction mismatch
Rewarded odour
Unreward-ed odour
(Relative to mismatch)
(Relative to next visual stimulus)
d
e
ΔR² movement
P-value odour prediction mismatch vs odour
P-value odour prediction mismatch vs no odour predicted
f
Cell response types
Type A
Type B
Mismatch
Non-significant
Running/licking neurons removed
3597 neurons
g
Striatal projecting cells
27%
4%
4%
65%
Non-striatal projecting cells
21%
8%
9%
61%

Extended Data Figure 2 | Response properties of ACC neurons. a. Mean responses of all non-striatal projecting ACC neurons (N = 4888 cells) aligned to task stimuli and behaviour. Activity is aligned to -1s to 3s around task stimulus or behaviour onset and mean baseline (-0.5s to 0s) is subtracted. Each condition is sorted by the averaged activity from 0 to 1s. **b.** Mean response of all negative mismatch neurons (N = 186 cells) aligned to the predicted odour mismatch event, and to the next visual stimulus onset (grey shading). Response shown to the predicted odour mismatch event (red) and actual odour delivery (green). Wilcoxon signed-rank test mismatch response vs baseline averaged 0 to 1.5 s from mismatch event and -0.5 to 0 s from the next stimulus onset, *** $P = 1.61 \times 10^{-28}$, $P = 6.57 \times 10^{-17}$ respectively, mismatch response vs odour response,*** $P = 2.97 \times 10^{-19}$, $P = 1.26 \times 10^{-6}$ respectively. **c**. Mean response of all positive mismatch neurons (N = 168 cells) aligned to predicted odour mismatch event, and to the next visual stimulus onsets (grey shading), shown with responses to the predicted odour mismatch event (red) and both the actual rewarded odour (green) and unrewarded odour (grey) delivery. Wilcoxon signed-rank test between the predicted odour mismatch response and unrewarded odour response averaged 0 to 1.5s from mismatch event and -0.5 to 0.s from the next stimulus onset, *** $P = 1.69 \times 10^{-28}$, $P = 5.45 \times 10^{-15}$. **d.** As in **a**, but for striatal projecting neurons (N = 421 cells). **e**. The influence of running and licking on each cell is inferred by modelling all stimuli, reward delivery, licking and running, and comparing the R^2 with and without the running and licking terms included. The change in R^2 is plotted in colour code for each cell (N = 6878 cells). Mismatch cells have significantly different activity between the prediction error vs odour, and prediction error vs no odour conditions (lines for $P < .05$ shown on each axis). Cells with stronger influence of running and licking (more negative ΔR^2) are not enriched in mismatch cells. **f**. Proportions of cells as in Fig. 2f, after removing cells which were influenced by running and licking (remaining N = 3597 cells). **G**. Proportions of cells as in Fig. 2h, after removing cells which were influenced by running (remaining N = 202 and 2437 striatal projecting and non-striatal projecting cells respectively, same colour legend as **f**).



Extended Data Figure 3 | Responses in visual to odour block transitions. a. Proportions of cells, as in Fig. 2f, for the visual to odour block transition. N = 463 Type A, 6.7%, 2063 Type B, 29.9%, and 432 mismatch cells, 6.2%, total 6878 cells, 10 mice. **b**. Mean response of all positive mismatch neurons (N = 432 cells) aligned to the odour prediction mismatch event (arrival of an unexpected odour). Response shown to the odour prediction mismatch event (red), actual odour delivery of odour 1 (green) and odour 2 (grey), and trials from the visual block where odour is neither predicted nor delivered (blue). **c**. As in **b**, for primary visual cortex (V1). N = 33 Type A, 1.5%, 69 Type B, 3.2%, and 104 mismatch cells, 4.8%, total 2138 cells, 4 mice. **d**. Venn diagram showing the overlap between mismatch cells in odour to visual block transitions and mismatch cells in visual to odour block transitions. The proportion of overlapping cells is significantly higher than chance (bootstrap test, $P < 0.01$.



Extended Data Figure 4 | ACC inhibition impairs task switching. a. Average number of trials per session required to switch from an odour block to a visual block, following bilateral infusions of saline or muscimol into the ACC. Saline infusions, 13 sessions, median ± IQR here and below, 6.25 ± 10,

muscimol infusions, 14 sessions, 29 ± 27.75, N = 4 mice. Wilcoxon rank-sum comparison P = 0.03. **b.** Average number of trials required in a session to switch from a visual block to an odour block, control 2.68 ± 2, continuous ACC silencing 2.33 ± 2.17, continuous PL silencing 2.75 ± 4.83. Wilcoxon signed-rank comparing control and continuous ACC silencing sessions P = 0.81, control and continuous PL silencing sessions P = 0.62, continuous ACC silencing and continuous PL silencing P = 0.91. **c.** RL model with belief state fit to the data, with and without ACC silenced. Number of trials taken by the agent to switch from odour to visual blocks increased, leading to fewer one-shot transitions. The model fit the data best when the amplitude of the mismatch signal was reduced by a factor of 0.22. **d**. Same as **c**. for the visual to odour block transitions. The model fit the data best when the amplitude of the mismatch signal was reduced by a factor of 0.95. **e.** Proportion of one-shot transitions from odour to visual block. Grey circles represent means for individual mice (N = 8), lines represent means ± STD across mice. Mean ± STD for control 0.43 ± 0.33, continuous ACC silencing 0.18 ± 0.13, one trial un-silencing 0.71 ± 0.32. Chi-squared tests of proportion comparing control and continuous ACC silencing P = 0.0003, control and one trial un-silencing P= 0.0002, continuous ACC silencing and one trial un-silencing P = 0.0002. **f.** Number of trials required to accurately discriminate odours (>80% accuracy over 30 trials) and ignore irrelevant visual stimuli (100% accuracy over 30 trials) in first task-switching session, for mice with and without continuous optogenetic ACC inhibition (N = 8 and 10 mice respectively), median ± IQR, trials to discriminate odours control 27.5 ± 62, continuous ACC silencing 60 ± 70.5, trials to ignore irrelevant visual stimuli control 142 ± 30, continuous ACC silencing 105 ± 43.3. Wilcoxon rank-sum comparison of trials to discriminate odours P = 0.35, trials to ignore irrelevant visual stimuli P = 0.11.