# 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, 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 that allows animals to update their knowledge of current rules or contexts. Detecting discrepancies between predicted and observed events is essential for this process. However, little is known about how the brain computes cognitive prediction-errors and whether neural prediction-error signals are causally related to task-switching behaviours. Here we trained mice to use a prediction-error to switch, in a single trial, between responding to the same stimuli using two distinct rules. Optogenetic silencing and un-silencing, together with widefield and two-photon calcium imaging revealed that the anterior cingulate cortex (ACC) was specifically required for this rapid task-switching, but only when it exhibited neural prediction-error signals. These prediction-error signals were projection-target dependent and were larger preceding successful behavioural transitions. An all-optical approach revealed a disinhibitory interneuron circuit required for successful prediction-error computation. These results reveal a circuit mechanism for computing prediction-errors and transitioning between distinct cognitive states.

**Introduction**

Animals need to rapidly update their behaviour to survive in a changing environment, and such behavioural flexibility is studied experimentally using task-switching paradigms. Task-switching involves shifting between distinct cognitive rules or contexts, allowing flexible behaviour adapted to changing environmental demands [1]. The framework of predictive processing [2–4] provides a simple yet powerful way of describing flexible task-switching behaviour using three stages. First, animals maintain a previously learned prediction of what they expect to happen in the world at any given instant, i.e., a 'model of the world'. Second, they may detect discrepancies between predicted and observed events, a 'prediction error'. Third, the prediction error guides the updating of their model of the world and their ongoing behaviour. While this account has widespread support across humans, monkeys and rodents [2,5,6], there is no clear link between the neural representations of prediction errors and their causal requirement in updating mental rules and subsequent behaviour.

There is evidence for cognitive prediction-error signalling in the prefrontal cortex (PFC) in tasks across humans, monkeys, and rodents [7–13]. The ACC in particular plays a central role in detecting the need for updating behavioural rules, and implementing the updated rules [7–10,14–25]. However, it is unclear how prediction-error signals are distributed across the cortex,

how these signals are organised with respect to the projection targets of cortical neurons, and to what extent the amplitude of prediction-error responses is important in subsequent behaviour. Crucially, it is unclear whether prediction-error signals have any causal influence on the subsequent updating of an animal's behavioural strategy. Finally, the inhibitory circuit basis for computing cognitive prediction errors is largely unknown.

Determining the neural basis of cognitive transitions is challenging because it requires experimental control over an animal's internal model of the world. Animals need to demonstrably hold one rule in mind and transition to another distinct rule on noticing a violation of a prediction based on the current rule. Although animals can be trained on tasks which change or reverse rules in a block-wise manner [12,26–33], these transitions typically involve several tens to hundreds of trials of intermediate performance, making it difficult to directly relate neural activity to an update of cognitive rules. One-shot block transitions provide a significant advantage in this regard. In one-shot block transitions [34,35], a single error leads to a complete and persistent switch between distinct cognitive states. Thus, repeated one-shot block transitions within a session would allow directly relating neural prediction-error responses with the updating of mental rules.

The identification of temporally well-defined cognitive prediction errors would allow addressing a key question: What neural circuit compares predictions and observations to compute prediction errors? Computing prediction errors requires inhibition[2], and distinct inhibitory cell classes with their specific connectivity patterns provide the basis for current models of prediction error circuits[36]. In particular, vasoactive intestinal peptide (VIP) expressing interneurons are crucial in shaping cortical activity[37–39] and provide unique computational opportunities through a disinhibitory motif involving somatostatin (SOM) expressing interneurons[40]. While VIP interneurons have been hypothesized to be a key player in prediction error computation[36], directly testing this theory is challenging, as it would require measuring neural prediction errors while the activity of VIP interneurons is perturbed in animals performing a cognitive task. Importantly, any neural perturbation should not lead to a change in the task-switching behaviour itself, as this would confound any interpretation of the neural circuit consequences of the perturbation.

Here we studied the neural basis of cognitive prediction errors during task-switching in the mouse neocortex. We trained mice to perform largely one-shot block transitions triggered by cognitive prediction errors. The prediction error was an absence of an expected stimulus, allowing us to better isolate cognitive factors from reward and stimulus evoked factors [41]. Using behavioural modelling, widefield calcium imaging, and optogenetic silencing, we generated a cortex-wide map of prediction-error signals and established a specific role for the ACC in task-switching behaviour. Using two-photon imaging of the ACC we identified a population of prediction-error neurons and, crucially, established that the duration of the prediction-error signal corresponded with the requirement of the ACC in task-switching at timescales within and across trials. We finally use all-optical methods to bi-directionally modulate VIP interneuron activity and establish that VIP interneurons play a specific and causal role in the computation of prediction errors in the ACC. These results provide direct evidence for a causal role of prediction-error signals in the ACC in task switching, and identify a key inhibitory interneuron class required for cognitive prediction error computation.

The novelty of our findings is twofold, as described in detail below: First, no previous study has established a causal link between an identified neural prediction error signal and behavioural transitions. Second, we use this approach to identify an inhibitory interneuron class required for the successful computation of prediction-error signals. This provides the first experimental support for an inhibitory model of prediction-error computation in any cognitive task.

## Results

We trained head-fixed mice to switch between blocks of discriminating two visual stimuli or discriminating two olfactory stimuli (Figure 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 (Figure 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). Visual and odour stimuli were never presented simultaneously (Figure 1B, Supplementary Figure 1).

### One-shot cognitive task-switching behaviour

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 the transitions from odour to visual blocks: mice noticed the absence of an expected odour stimulus (an odour prediction error) to switch their behaviour, and accurately responded to the now-relevant visual stimuli in subsequent trials (Figure 1D-E). Thus, we focused on the transition from odour to visual blocks, triggered by the omission of an expected odour stimulus (but see below for visual to odour block transitions).

Animals typically require many trials to switch between blocks of distinct rules, transitioning through periods of intermediate performance [12,26–32,42]. This makes it difficult to precisely identify the behavioural and neural processes underlying a rule-switch. Optimal 'one-shot' transitions offer a significant advantage, since they involve animals switching between two accurately applied rules after a single error [34,35]. Most odour to visual block transitions in our task were indeed one-shot (Figure 1D-F, 63.2% one-shot transitions, see Supplementary Movie 1 for an example). We used strict criteria to ensure that the one-shot block transitions captured a complete mental transition between two demonstrably applied rules (see methods). A low probability of licking in response to irrelevant visual stimuli in the odour block led to a corresponding proportion of zero-shot block switches, or 'fluke transitions' (Figure 1F, zero trial bin) which were removed from further analysis. The remaining transitions were dominated by one-shot transitions (Figure 1F arrowhead) with a smaller fraction requiring two or more error trials.

To understand the processes underlying this rapid task-switching, we first fitted the behaviour to a basic reinforcement learning (RL) model (tabular SARSA, Figure 1G, top) in which an agent

continually 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 (Figure 1H, top), and the best fit of learning and exploration rates produced slower block transitions (Figure 1H, middle, see methods). We next built an RL model with a context belief state, hierarchically controlling subsequent choices (Figure 1G, bottom). 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, capturing the two rules of the task. Once the learning was complete, the agent was able to rapidly switch between these two Q-value-tables on accumulating a large enough context-error signal. This belief state model reproduced the rapid block transitions (Figure 1H, bottom). Once the model had learnt the task, freezing all learning, i.e., Q-value updates, did not affect the block transitions (see methods), demonstrating that switching between blocks did not involve any further learning. Instead, the agent inferred the block transitions using the context error. These results suggest that rapid task-switching behaviour involves transitioning between abstract representations of rules or contexts, driven by prediction errors.

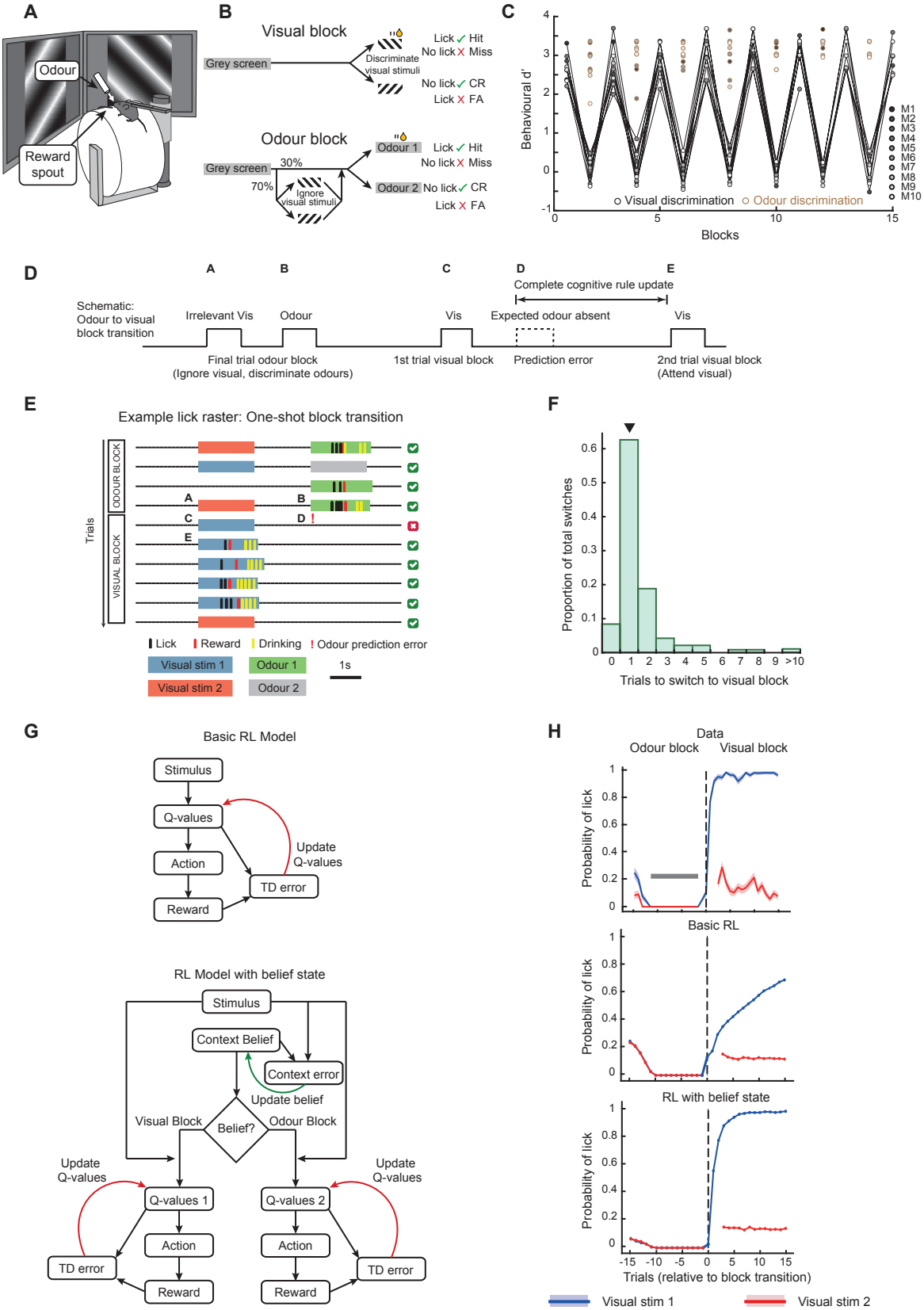



**Figure 1 | Mice perform rapid task-switching with one-shot block transitions. A.** Experimental setup. **B**. Task schematic. Mice switched between blocks of discriminating two visual stimuli in the visual blocks or discriminating two olfactory stimuli while ignoring the same visual stimuli in the 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.** Schematic of a transition from an odour block to a visual block, showing the last trial of the odour block and first two trials of the visual block, indicating the moment of odour prediction error. Labels A-E refer to timepoints in E. During one-shot block transitions as in E, a complete cognitive rule update occurs in the time indicated. **E.** Example behaviour 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 started accurately discriminating the same visual stimuli. **F.** Histogram showing the number of trials required to switch between the two blocks, with one-shot transitions indicated by an arrowhead. N = 95 transitions, 17 sessions, 14 mice. **G**. Schematic of reinforcement learning (RL) models. Top, basic RL, bottom, RL with belief state. In the RL model with belief state, Q-values were only updated while the agent learnt the task, following which only context belief was updated to switch between blocks. **H**. 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 95 odour to visual block transitions. Shading indicates SEM. Grey bar indicates the 10-trial duration where mice performed at 100% accuracy with no licks to either irrelevant visual stimulus, which was a condition for triggering a block transition. First 3 trials of each block were forced to be Visual Stimulus 1 (to assess the animal's belief of which block it was in), resulting in a gap in the red curve depicting Visual Stimulus 2. Middle, basic RL model fit to data. Bottom, RL model with belief state fit to data.

## Map of prediction error signals

Which brain regions signal the prediction error at block transitions? We measured neural activity across the entire dorsal cortical surface using widefield calcium imaging (Figure 2A) to identify the cortical regions which represent cognitive prediction-error signals at block transitions. Our task design provides a well-defined moment of prediction error at the odour to visual block transition: when, following the visual stimulus, an expected odour does not arrive (Figure 1D). Any neural activity specific to this prediction error 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 (the prediction error), and (3) absent once again when the odour was no longer expected later in the visual block (schematic of the three conditions shown in Figure 2B).

We aligned the hemodynamic-corrected and movement-corrected calcium activity to these three temporal epochs of odour onset, odour prediction error, and no odour predicted (Figure 2b). The activity in these epochs evolved across the cortex over time, with higher activity apparent in the odour prediction-error condition (Figure 2C, first three columns). We identified pixels at each time point which satisfied the criteria described above (see methods) and mapped their activity (Figure 2C final column, colour map represents t-statistic of comparison between first and second column). This revealed the precise spatio-temporal evolution of the prediction-error signal across dorsal cortex, which spread from anterior secondary motor regions to posterior parietal areas (average map from 1.2 to 2.7s shown in Figure 2D). In particular, we found a strong prediction-error signal in areas superior to prefrontal cortex. Although the visual cortex and olfactory bulb showed robust sensory stimulus evoked responses to visual and odour stimuli respectively, they did not contain

192 detectable prediction-error signals (Figure 2D, averaged activity PSTHs shown in Figure 2E).
193 Together these results suggested that prefrontal cortical areas, not primary sensory areas,
194 are involved in detecting the cognitive prediction error relevant for the rapid block transitions
195 in this task.

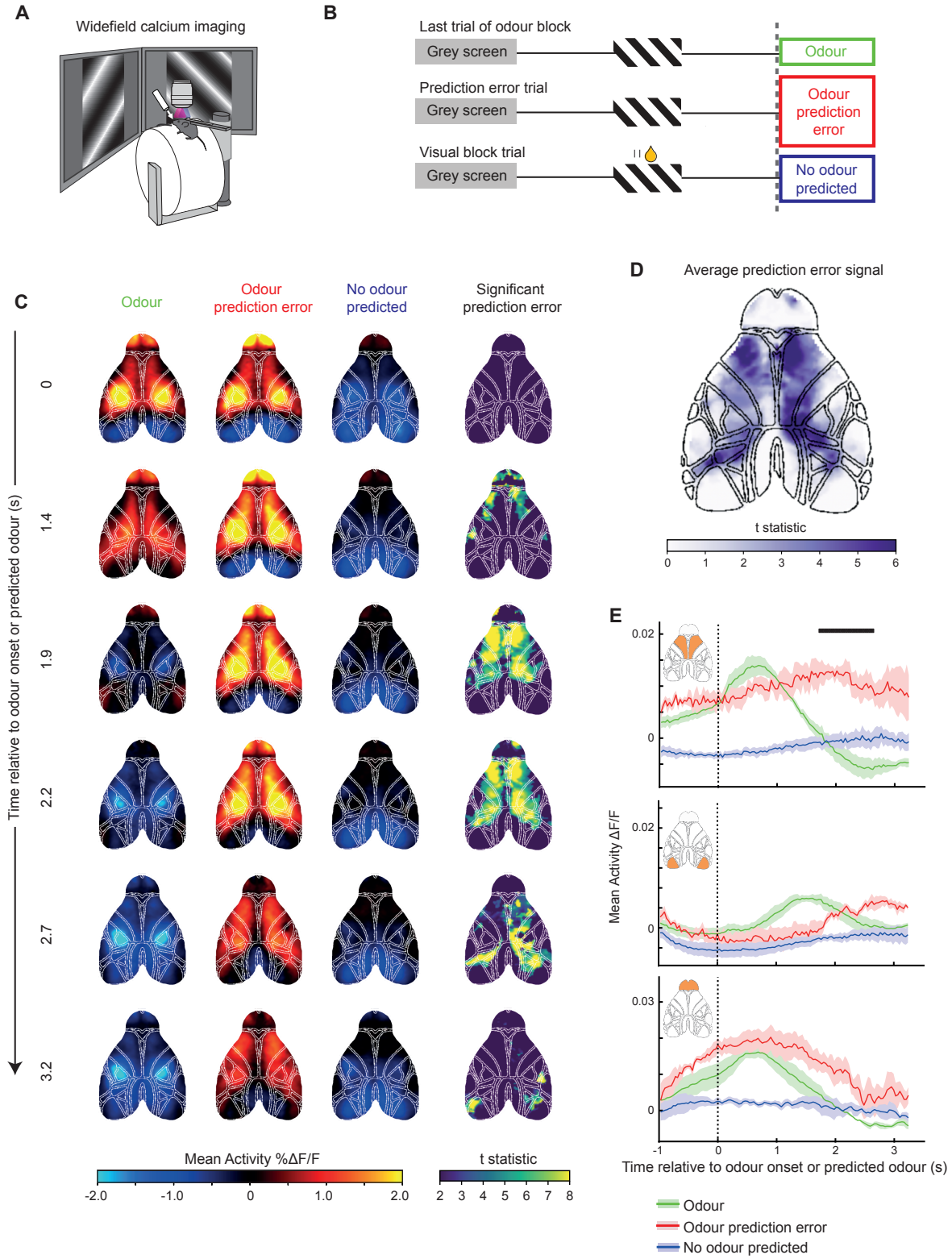



**Figure 2 | Cortex-wide map of prediction-error signalling. A.** Schematic of widefield calcium imaging. **B.** Schematic of behaviour during a one-shot transition from an odour to visual block, indicating the three trial types used to identify prediction-error pixels. The first trial of the visual block contains an odour prediction error, due to the absence of an expected odour stimulus, which mice can use to infer the change in block type. Prediction-error pixels were defined as pixels with significantly different

activity between the odour prediction-error period and both the odour period and no odour predicted period. **C**. 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 error and odour trials, and between odour prediction error and no odour predicted trials, with the t-statistic values from the comparison of column 1 and 2 displayed in colour code. **D.** Pixels with significant prediction-error signalling, averaged across 1.2-2.7s relative to the expected/actual odour onset. **E.** Activity profiles from the secondary motor cortex (top), primary visual cortex (middle) and olfactory bulb (bottom, ROIs shown in insets). Black line indicates periods during which the odour prediction-error response was significantly different from both the odour and no odour conditions (paired t-test, $P < 0.01$ for more than 500ms, N = 4 mice).

## ACC is required for rapid block transitions

We next asked which specific region within prefrontal cortex was required for driving the task-switching behaviour. A number of studies have linked ACC activity to surprising events [9,42], errors [10,23] or negative feedback [43,44], which promote updating of behavioural policies [16,21], or promote control [15]. In addition, the prelimbic (PL) cortex has been implicated in task-switching behaviour [45]. If prediction-error signalling in the ACC or PL is required for rapid task-switching, silencing these regions should disrupt the behaviour specifically at block transitions. To test this, we optogenetically silenced the ACC or PL bilaterally for the entire behavioural session in a group of mice (Figure 3A). ACC silencing caused a strong deficit in switching from odour to visual blocks, when the mice needed to detect the absence of an expected odour to switch rules. Strikingly, with the ACC silenced, mice ignored the rewarded visual stimulus repeatedly in anticipation of an odour stimulus, reflecting the continuing application of the odour block rule (Figure 3B). As a result, ACC silencing increased the number of trials taken to switch (Figure 3C, D) and reduced the proportion of one-shot odour to visual block transitions (Figure 3D, E). This was also the case when silencing the ACC using bilateral infusions of the GABA-A receptor agonist muscimol (Supplementary Figure 2C). The belief-state RL model could fit the ACC silencing data with the prediction-error signal reduced by a factor of 0.22, suggesting that the task-switching deficit could be explained by a partial reduction in the ACC prediction-error signal (Supplementary Figure 2D). These results demonstrate that the ACC is required for processing the omission of an expected event, a prediction error, during task-switching.

Interestingly, silencing PL in the same mice did not affect switching behaviour (Figure 3C). Although we observed a cognitive prediction-error signal across a substantial portion of frontal cortex in our widefield experiment (Figure 2D), the absence of any behavioural effect of silencing PL, adjacent to the ACC, reveals a striking specificity in the role of the ACC. These results indicate that the ACC has a specific role in rapidly updating behavioural rules or contexts within seconds, and this role is not widely shared with other prefrontal areas.

While the ACC is required for rapidly switching between task rules, does its role extend to applying these task rules once the animal has switched blocks? We found that even during continuous ACC silencing, once the mice did spontaneously switch from ignoring to discriminating the visual stimuli, the subsequent accuracy of visual discrimination was only slightly lower from unsilenced sessions (Figure 3F). This suggests that the ACC plays a critical role in transitioning between task rules, and less so in maintaining these rules.

In addition to rapid task-switching, the ACC has been implicated in guiding slower learning processes [46–48]. We optogenetically silenced the ACC in a subset of mice as they first learned the switching task and found no difference relative to controls in either the rate of learning a novel stimulus-reward association, or learning to ignore irrelevant stimuli [49] (Supplementary Figure 2G). Thus, in this paradigm, the role of the ACC is specific to rapid task-switching and does not extend to slower learning.

Overall, these results established that the ACC is essential for rapid task switching driven by a cognitive prediction error. However, ACC projections become dispensable when task demands are reduced in an attentional task [50]. We asked if a similar result was true in task-switching behaviours, that is, if an animal could overcome the inhibition of the ACC if the block transition was marked by a more salient event than the absent expected odour. The opposite direction of block transition in our task, from visual to odour blocks, was marked by the unexpected arrival of an odour, a more salient prediction error. Most visual to odour block transitions (60%) were also one-shot (Supplementary Figure 3A, B) and were captured well by the belief-state RL model (Supplementary Figure 3C). Interestingly, silencing the ACC had no effect on the switching behaviour in these visual to odour block transitions (Supplementary Figure 3D-F). Thus, when task demands are reduced by using a highly salient event to mark block transitions, the ACC may become dispensable in task switching, similar to attentional tasks. Our study, however, focused on the odour to visual block transitions, which relied on the ACC.

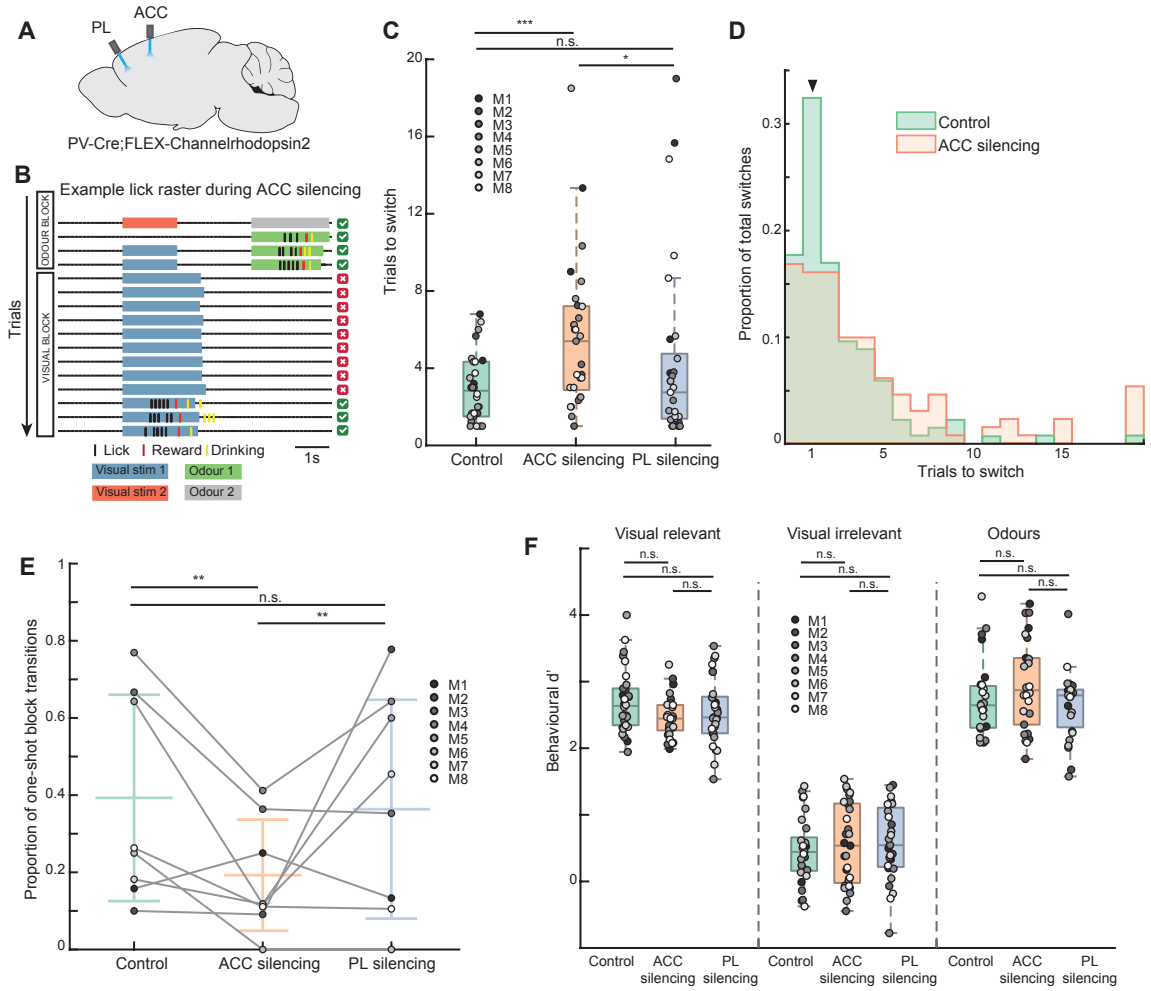



**Figure 3 | Optogenetic inhibition of ACC impairs task-switching. A.** Schematic of bilateral optic fibre implantation targeting both ACC and PL in each mouse. **B**. Example lick raster showing stimuli, lick and reward times during a transition from an odour to visual block during ACC silencing leading to impaired switching behaviour. **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 PRL silencing P = 0.78, continuous ACC silencing and continuous PRL silencing P = 0.02. **D.** Histogram of number of trials taken to switch from odour to visual blocks during control and continuous ACC silencing sessions. N = 136 and 131 control and ACC silencing transitions respectively. One-shot transitions indicated by arrowhead. **E.** Proportion of transitions from odour to visual blocks that were 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.** Steady-state stimulus 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.

289

290 **Prediction-error signals in ACC neurons**

291 Since block transitions were typically complete within a single trial, this also meant that the
292 neural activity in the ACC required for the block transitions must be present within the few
293 seconds between the prediction error (absent predicted odour) and the next visual stimulus
294 (Figure 1D). We therefore asked what neural signals in the ACC during this period may account
295 for its role in rapid block transitions. We recorded the activity of populations of ACC neurons
296 during the behaviour using chronic two-photon calcium imaging through a microprism (Figure
297 4A, Supplementary Figure 4A). As expected, we found that individual ACC neurons responded
298 to many task-related variables. Subsets of neurons responded to visual and odour stimuli,
299 locomotion onsets, reward delivery and licking (Figure 4B, C). A binary classifier was able to
300 accurately decode the identity of visual and odour stimuli after stimulus onsets from ACC
301 neural population activity, and was further able to decode the block type both before and
302 after stimulus onsets (Figure 4D). Thus, ACC contained diverse signals relevant to the ongoing
303 task.

304 To identify neurons representing prediction errors, we asked if a neuron showed responses
305 which were (1) absent in the final trials of the odour block (2) present during the odour
306 prediction error, and (3) absent once again when the odour was no longer expected later in
307 the visual block (Figure 4E, the same criteria as used in the widefield data). We found neurons
308 in the ACC that satisfied these criteria, which we termed 'prediction-error neurons'. Figure 4F
309 shows an example prediction-error neuron which was suppressed in trials when the expected
310 odour did arrive, but responded strongly during the odour prediction error (when an odour
311 was expected but did not arrive), and was not responsive later when an odour was no longer
312 expected (see also Supplementary Figure 4B). Importantly, the prediction-error signal is a
313 response to the non-occurrence of an expected stimulus, not a stimulus onset or offset,
314 ensuring that the response originates from a violated stimulus expectation.

315 The average of all prediction-error neurons positively activated by the prediction error is
316 shown in Figure 4G. The percentage of all recorded neurons which showed any prediction-
317 error responses was 9% (Figure 4H). This group contained neurons with a positively activated
318 response to the prediction error (Figure 4F, g, 27% of prediction-error neurons), as well as a
319 similar number of neurons with an inhibited response to the prediction error (30% of
320 prediction-error neurons, Supplementary Figure 5A) and other combinations of activity
321 profiles. Response inhibition by the arrival of the expected odour occurred for both rewarded
322 and non-rewarded odours, suggesting that these neurons did not reflect a simple negative
323 reward signal (Supplementary Figure 5B). Crucially, changes in running and licking could not
324 account for the identification of prediction-error neurons (Supplementary Figure 5C, D).

325 To investigate whether prediction-error neurons were found widely across the brain, we
326 conducted the same experiment using two-photon calcium imaging to record activity in the
327 primary visual cortex (V1) (Figure 4I, left). Consistent with the widefield calcium imaging
328 results, our findings revealed a near absence of prediction-error neurons in V1 (Figure 4I,
329 right). This corroborated the widefield imaging results and, importantly, confirmed that our
330 identification of prediction-error neurons in the ACC was statistically reliable.

331 We also studied the opposite direction of block transition, from visual to odour blocks, which
332 was marked by the unexpected arrival of an odour stimulus. We again found prediction-error

neurons in the ACC which responded differently to the unexpected odour, compared to the same odour when expected, or no odour (Supplementary Figure 6A, B, similar proportions of prediction-error neurons were obtained when controlling for running and licking behaviours, data not shown). Crucially, however, we obtained a similar proportion of these neurons in V1 (Supplementary Figure 6C). Thus, prediction-error neurons signalling an unexpected appearance of an odour are found both within and outside prefrontal cortex, consistent with the finding that the ACC is not necessary for this direction of transition (Supplementary Figure 3d-f). Interestingly, neurons which represented both directions of prediction error were present at a higher proportion than expected by chance (Supplementary Figure 6D), suggesting the presence of generalist, in addition to specialist prediction-error neurons in the ACC.

**Striatal projections in the ACC exclude prediction errors**

Prediction-error neurons may broadcast their activity widely across the brain, or they may be enriched or excluded from populations based on their projection target. We distinguished between these two scenarios 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 [51]. CTB-Alexa647 injections in the striatum identified striatal-projection neurons in the ACC (Figure 4J), and non-retrolabelled neurons were enriched for non-striatal projecting neurons. Although striatal projecting neurons had mostly overlapping response properties with non-striatal projection neurons (Figure 4K, Supplementary Figure 5E, F), they contained a significantly smaller proportion of prediction-error neurons (Figure 4L, Chi-squared test of proportion P = 0.0015, striatal projecting prediction-error neurons = 21/410, 5%, non-striatal projecting prediction-error neurons = 463/4888, 9%). This lower proportion of striatal projecting prediction-error neurons was outside the 99% confidence intervals of the non-striatal projecting prediction-error neurons (Figure 4M, bootstrap test). Thus, prediction-error responses in the ACC are not indiscriminately broadcast, but are significantly excluded from a major projection target.

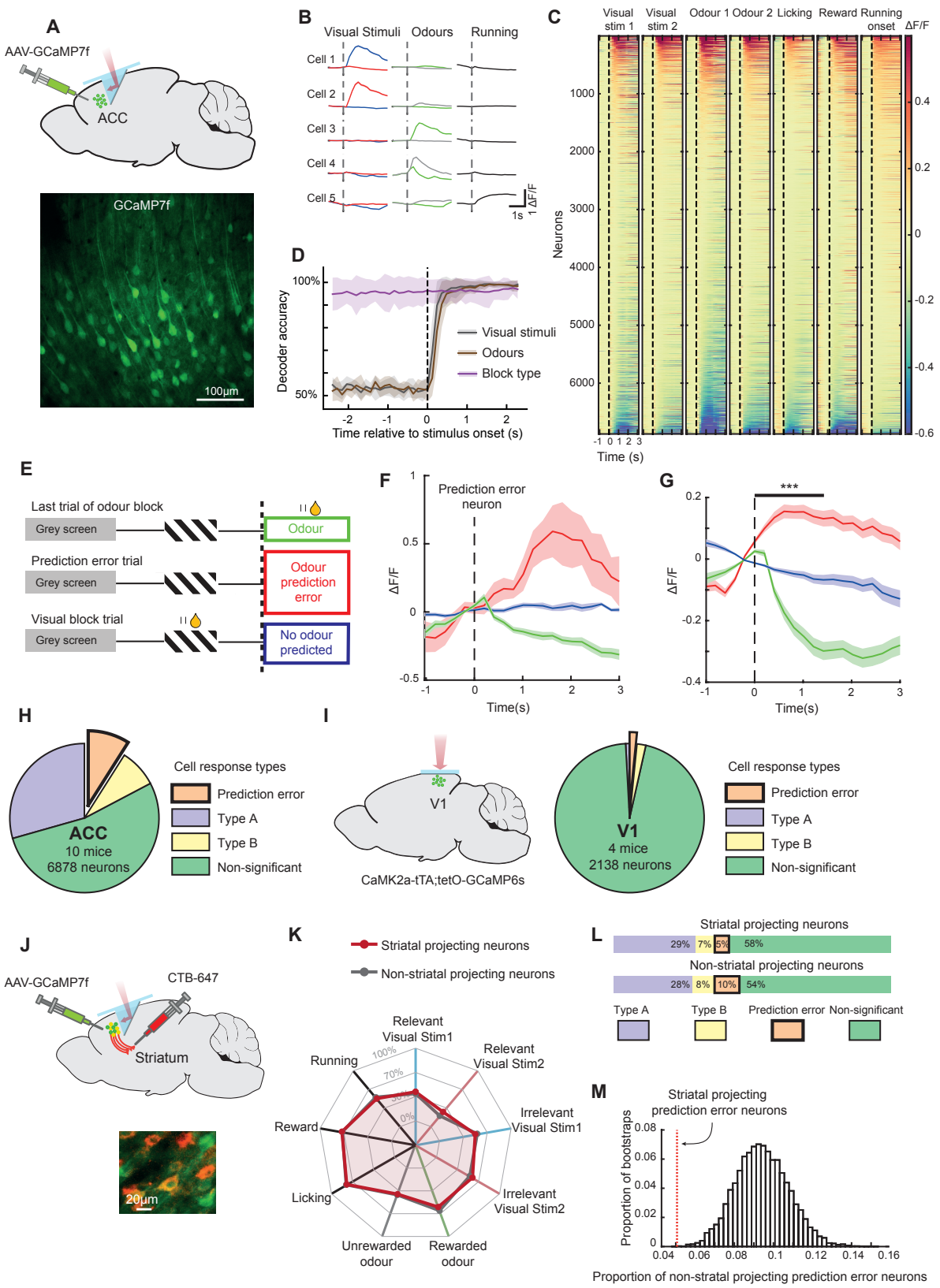



**Figure 4 | ACC neurons encode prediction-error signals. A.** Top, schematic depicting two-photon imaging of neurons expressing GCaMP7f in the ACC through a microprism implant. Bottom, example ACC imaging site. **B.** Responses from 5 example neurons to task stimuli and running onsets. Shading

here and below indicates SEM. **C**. Mean responses of all ACC neurons (N = 6878 neurons) aligned to task stimuli and behaviour. Activity was aligned −1s to 3s around task stimulus or behaviour onset and mean baseline (−0.5s to 0s) subtracted. Each condition is sorted by the averaged activity from 0 to 1s. **D**. Time course of decoding accuracy of a binary classifier using neuronal activity of simultaneously imaged neurons 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. **E**. Schematic of behaviour during a one-shot transition from an odour to visual block, indicating the three trial types used to identify prediction-error neurons. **F**. Example prediction-error neuron with a significantly larger response to the odour prediction-error condition (red), compared to the actual delivery of the odour (green) or trials where odours were neither predicted nor delivered (blue). **G**. Average response of all prediction-error neurons with a positive response to the odour prediction-error condition, in the three conditions as described in **F** (N = 168 neurons, Wilcoxon signed-rank test between the prediction error and odour conditions averaged 0 to 1.5s, *** $P = 1.69\times10^{-28}$). **H.** Proportions of neurons with significantly different activity (averaged 0 to 1.5s) between the three trial types described in **e**. Type A and Type B were significantly different only when comparing odour prediction error to odour (N=2007 neurons, 29%) or to no odour predicted conditions (N = 567 neurons, 8%) respectively. Prediction-error neurons were defined as neurons significantly different in both comparisons (N = 616 neurons, 9%, total 6878 neurons, 10 mice). **I.** Left, schematic of two-photon calcium imaging from primary visual cortex (V1). Right, same comparisons as in **H** with neurons recorded from V1 (N = 18 Type A, 1%, 41 Type B, 2%, and 35 prediction-error neurons, 1.6%, total 2138 neurons, 4 mice) **J**. Top, schematic of retrograde labelling and imaging strategy. Bottom, example image of retrogradely labelled striatal-projecting (CTB-Alexa647 labelled) and non-striatal projecting neurons in the ACC. **K**. Percentages of recorded neurons significantly responsive to each of 9 task events. **L**. Proportion of prediction-error neurons from striatal-projecting and non-striatal projecting neurons in ACC. Striatal projecting: N =124 Type A (29%), 30 Type B (7%), 21 prediction-error neurons (5%), total 421 neurons. Non-striatal projecting: N = 1384 Type A (28%), 392 Type B (8%), 484 prediction-error neurons (10%), total 4888 neurons, 8 mice). **M**. Bootstrapped distribution of proportion of prediction-error neurons in non-striatal projecting neurons. Data for proportion of striatal projecting prediction-error neurons indicated as vertical line. Data is outside the 99% confidence intervals.

## Prediction-error signals in the ACC coincide with effective silencing and un-silencing epochs

Having identified prediction-error neurons in the ACC, we next asked whether the duration of the neural prediction-error signal corresponded to the duration of causal involvement of the ACC in the task. We first asked how sustained the prediction-error response was within an individual trial. We focused on all ACC neurons with a positive prediction-error response and found that their response peaked soon after the time of expected odour onset, but remained significantly higher than both its baseline and the actual odour response until the beginning of the next trial (Figure 5A, Wilcoxon signed-rank test prediction-error response vs baseline aligned to prediction error and next stimulus $P = 2.98\text{x}10^{-15}$, $P = 3.34\text{x}10^{-8}$ respectively, prediction-error response vs odour response, $P = 1.69\text{x}10^{-28}$, $P = 5.45\text{x}10^{-15}$ respectively). Indeed, this corresponded to the duration of the requirement of the ACC in the task, since silencing the ACC both during the ITI and peri-stimulus period on each trial (Figure 5B) caused significant deficits in the switching behaviour, although each to a smaller degree than continuous silencing (Figure 3C). Thus, ACC activity is required not only around the moment of expectation violation, when the largest prediction-error responses are present,

but remains important until the start of the next trial as the animal updates its belief about the current rule.

Is the ACC required for multiple trials as a block transition occurs? To address this question, we first asked how sustained the prediction-error response in the ACC was across consecutive trials. We found that the prediction-error response was present only on the first prediction-error trial, and rapidly decayed to non-significant amplitudes on subsequent trials (Figure 5D). If the ACC indeed enables block transitions through prediction-error signalling, the rapid decay of the prediction-error signal over trials would predict that the ACC is only required during the prediction-error trial, and not earlier or later. To directly test this prediction, 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 (Figure 5E). The resulting behaviour showed no deficits in task switching, and mice rapidly switched between blocks (Figure 5F, bottom, Supplementary Figure 2E). Interestingly, the speed of behavioural switching in the un-silencing condition was significantly faster than controls, possibly due to enhanced ACC activity from the removal of inhibition (see also Figure 6). Critically, after switching to the new block, the mice performed highly accurate discrimination for the rest of the block despite the ACC being continuously silenced (Supplementary Figure 2F). The rapid task switching could not be accounted for by a startle response to the light offset, since the mice did not lick in response to the light offset, or the visual stimulus immediately following the light offset (middle row in Figure 5E), but instead licked in response to the subsequent visual stimulus, after the odour prediction error. Thus, the ACC is specifically required only for processing the prediction error and is thereafter no longer required for accurate task performance. These results provide direct evidence for prediction-error signals in the ACC driving task-switching [2,4], and are consistent with a prominent theory suggesting that the PFC largely signals the non-occurrence of expected outcomes [52].

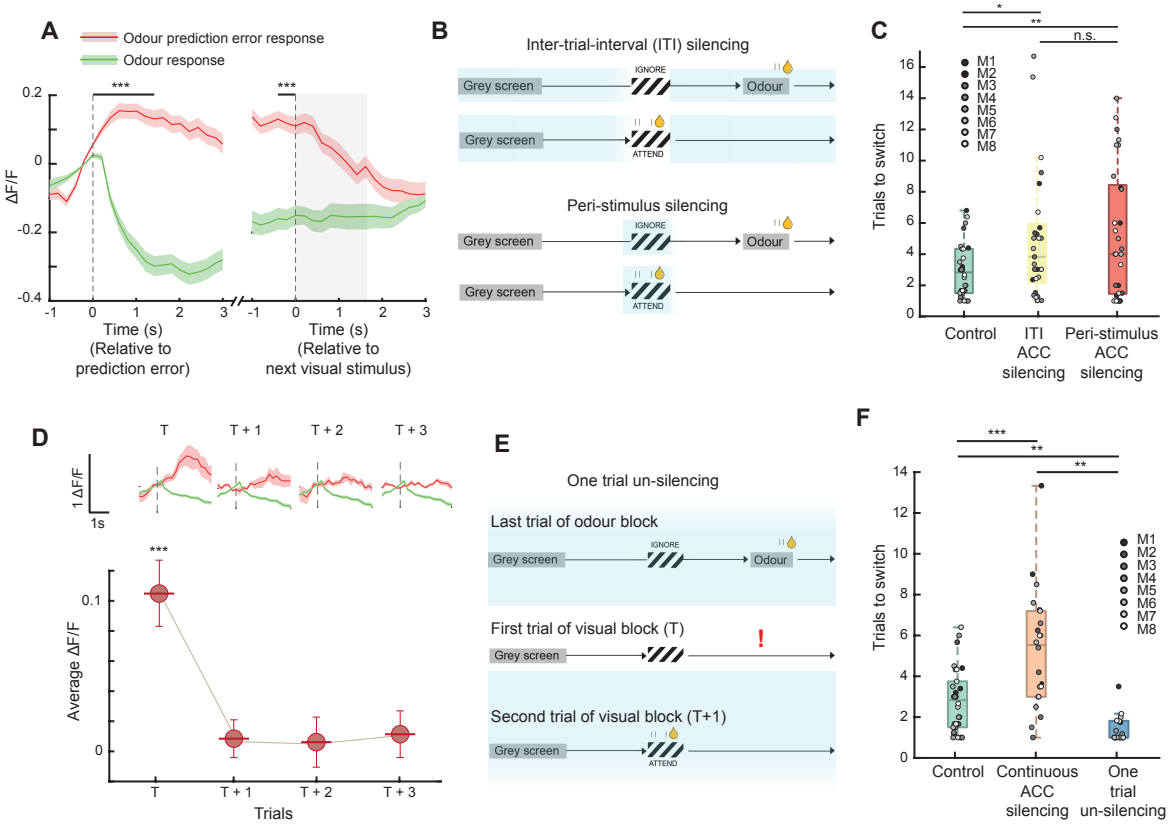



**Figure 5 | ACC prediction-error signal duration corresponds to requirement of the ACC in task switching. A.** Mean response of all positively responding prediction-error neurons (N = 168 neurons) aligned to the odour prediction-error event, and to the next visual stimulus onset (grey shading). Response shown to the odour prediction-error event (red) and actual odour delivery (green). Wilcoxon signed-rank test between the two conditions averaged 0 to 1.5s from prediction-error 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. **B**. Schematics depicting inter-trial-interval (ITI) and peri-stimulus silencing epochs. **C**. Number of trials required in a session to switch from an odour to visual block, median ± IQR here and below, 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. **D.** Mean odour prediction-error response relative to pre-event baseline aligned to the same temporal epoch over consecutive trials (T is first prediction-error event in an odour to visual block transition), shown up to 3 trials following first prediction-error 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 neuron responses showing the rapid decay of prediction-error response over trials. **E**. Schematic depicting silencing during the entire session except for one trial at each odour to visual block transition. **F**. Number of trials required in a session to switch from an odour to visual block, 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.

468 **Larger prediction-error signals in ACC precede successful one-shot block transitions**

469 Finally, to determine whether prediction-error neurons in the ACC actively contribute to the
470 behavioural transitions across blocks, we asked if the amplitude of prediction-error responses
471 at block transitions was related to the subsequent success in behavioural switching. We
472 divided the odour to visual block transitions into one-shot and slower than one-shot block
473 transitions (Figure 6A, N = 51% one-shot and 49% slower transitions). Prediction-error
474 neurons with a positive response were more active preceding one-shot block transitions,
475 compared to slower transitions (Figure 6B) when aligned to the prediction-error event
476 (Wilcoxon signed-rank test, P = 0.0003) and also when aligned to the next stimulus onset (P =
477 0.019, N = 146 neurons). This result was also true when combining all prediction-error
478 neurons ($P = 8.57 \times 10^{-9}$, N = 616 neurons). This result could not be accounted for by differences
479 in running speed across the two conditions (Supplementary Figure 7A). The pupil diameter
480 also did not show significant differences between the one-shot and slower block transitions,
481 and largely reflected running speed changes (Supplementary Figure 7B).

482 Since our belief-state RL model of the behaviour included an explicit context-prediction-error
483 signal, we asked if this signal also predicted one-shot block transitions across contexts. The
484 model revealed a similar pattern, where the amplitude of the noisy prediction-error signal
485 was predictive of future one-shot transitions (Figure 6C, Wilcoxon rank sum test, $P < 10^{-8}$, N =
486 70 transitions). These results support the claim that prediction-error neurons in the ACC play
487 a key role in driving rapid task-switching behaviour, particularly when they produce larger
488 prediction-error signals.

489

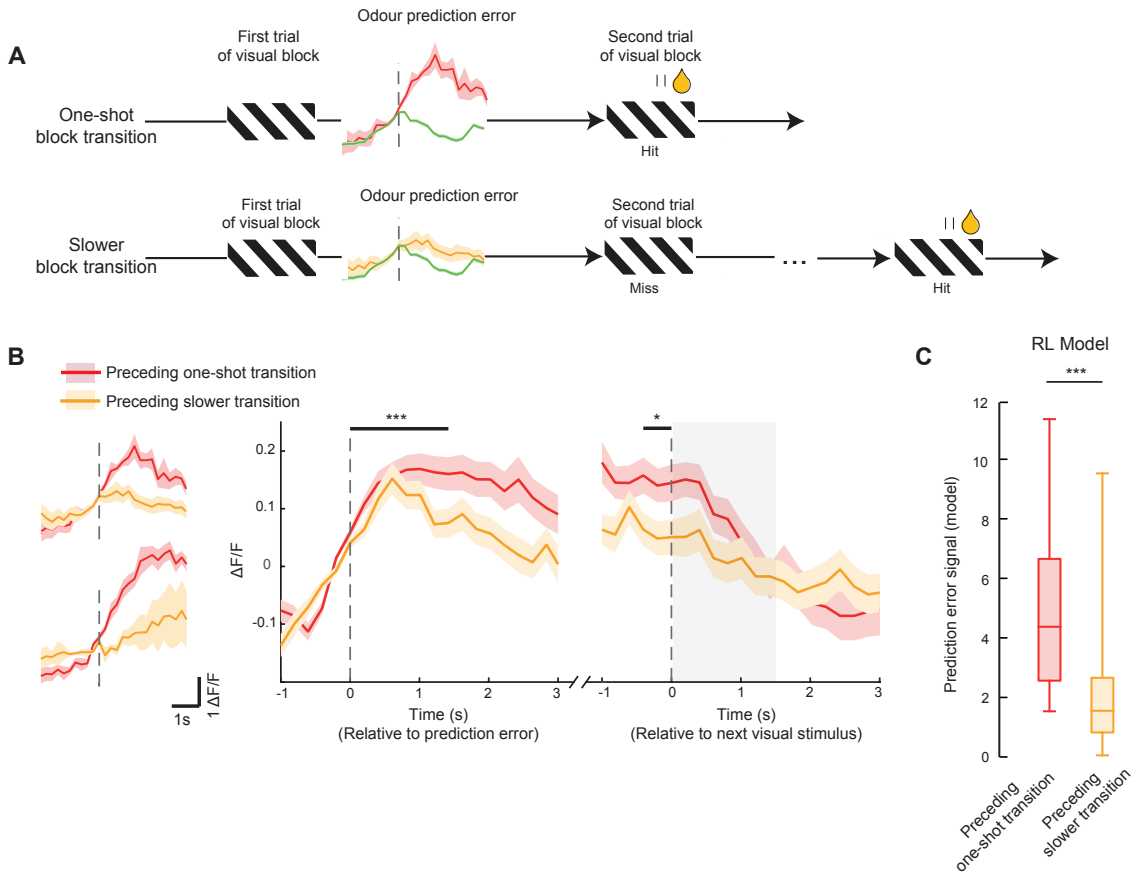



**Figure 6 | Larger prediction-error signals in ACC precede one-shot block transition behaviour. A.** Schematic showing odour to visual block transitions and the prediction-error response preceding two scenarios, top, where the mouse executes a one-shot block transition, or switch, and bottom, where the block transition is slower than one-shot. Red and yellow responses are prediction-error responses from an example neuron, green shows the response to delivery of odour, shading indicates SEM here and below. **B**. Left, odour prediction-error response of two example neurons preceding a one-shot (red) or slower transition (yellow). Right, mean responses across all positively-responding prediction-error neurons (N = 146 neurons) preceding one-shot and slower transitions (Wilcoxon signed rank test between the two conditions averaged 0 to 1.5 s from prediction-error event and -0.5 to 0 s from the next stimulus onset, *** P = 0.0003, * P = 0.019). **C**. Prediction-error signal amplitude from the Belief-State RL model, preceding one-shot transitions and slower transitions, *** $P < 10^{-8}$, N = 70 transitions.

## VIP interneurons contribute to prediction-error computation in the ACC

Local inhibitory circuits are crucial in shaping cortical activity[53], and are believed to be necessary for computing prediction errors[2]. These circuits need to compare predictions and observations, with a mismatch between the two leading to prediction-error signals. While a diversity of inhibitory circuits may compute prediction-errors,[36] VIP interneuron disinhibition[37–39] is hypothesized to be key in this process[36] (Fig 7A). Indeed, VIP driven disinhibition interacts with a perceptual prediction error during visuo-motor coupling mismatch in V1[54]. We asked whether VIP interneurons played a role in producing cognitive

prediction-error responses in the ACC in our task. We employed an all-optical approach, where we photoactivated or photoinhibited VIP interneurons while simultaneously measuring the activity of VIP and non-VIP neurons in the same brain region with in-vivo two-photon calcium imaging (Figure 7A, B). Importantly, the VIP perturbations were in one hemisphere only, and did not lead to any changes in the task-switching behaviour itself (Wilcoxon signed-rank test comparing behavioural d' for the relevant visual stimuli, irrelevant visual stimuli, and odours between control and ITI VIP activation or inhibition sessions, all Ps > 0.05, similar Wilcoxon signed-rank test comparing the number of trials to switch between blocks all Ps > 0.05).

On photoactivating VIP cells in the ACC in passive mice with increasing light powers (Figure 7C bottom, inset), we observed monotonically increasing population activity in non-VIP cells (Figure 7C). Since non-VIP cells in cortex are predominantly excitatory pyramidal neurons, this experiment demonstrated the effective disinhibition induced by our all-optical approach. We next photoactivated VIP cells in the ACC as mice performed the task-switching behaviour. VIP cells were photoactivated on each trial only during the inter-trial-interval (ITI) period, which encompassed the prediction-error event at block transitions (Figure 7D). We compared the proportion of prediction-error neurons in the VIP photoactivation sessions to control sessions and found that VIP activation during the ITI period strongly reduced the percentage of prediction-error neurons from 10% to 2% (Figure 7E, Chi-squared test of proportion P < 0.0001). To determine how VIP photoactivation affected the response of ACC neurons to prediction-errors, we measured the average response amplitude to the odour prediction-error event, and odour stimulus, from positively responding prediction-error neurons (Figure 7F). We found that VIP photoactivation led to a significant reduction only in the response to the odour prediction-error, demonstrating that VIP photoactivation limits the prediction-error response amplitude (Figure 7F bottom). Critically, we confirmed that our VIP activation did not lead to widespread deficits in ACC responses, since our perturbation did not affect representations of stimuli and block rules. A binary classifier as used in Figure 4D 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 (Figure 7G) and during the ITI period (data not shown). Thus, inducing disinhibition through VIP interneuron activation specifically and heavily disrupted prediction error signalling in the ACC.

Finally, to prove their necessity in generating prediction error responses in the ACC, we optogenetically silenced VIP interneurons during the task. VIP inhibition during the ITI period strongly reduced the percentage of prediction-error neurons from 5% to 0.6% (Figure 7H, Chi-squared test of proportion P < 0.0001). Again, a binary classifier was able to accurately decode visual and odour stimuli and block type, both when VIP cells were photoinhibited during the visual stimulus onsets (Figure 7I) and during the ITI period (data not shown). Crucially, light-only control mice which did not express any opsin showed no change in the proportion of prediction-error neurons detected on presentation of the same optogenetic light (Figure 7J, 3.6% to 4.3%, Chi-squared test of proportion P = 0.46). Together, these results demonstrate that bidirectional perturbation of VIP interneurons strongly and specifically disrupts prediction-error computations in the ACC. Future work may test the role of other cell classes such as PV interneurons, and the sources of prediction and observation signals in this computation (Figure 7A).

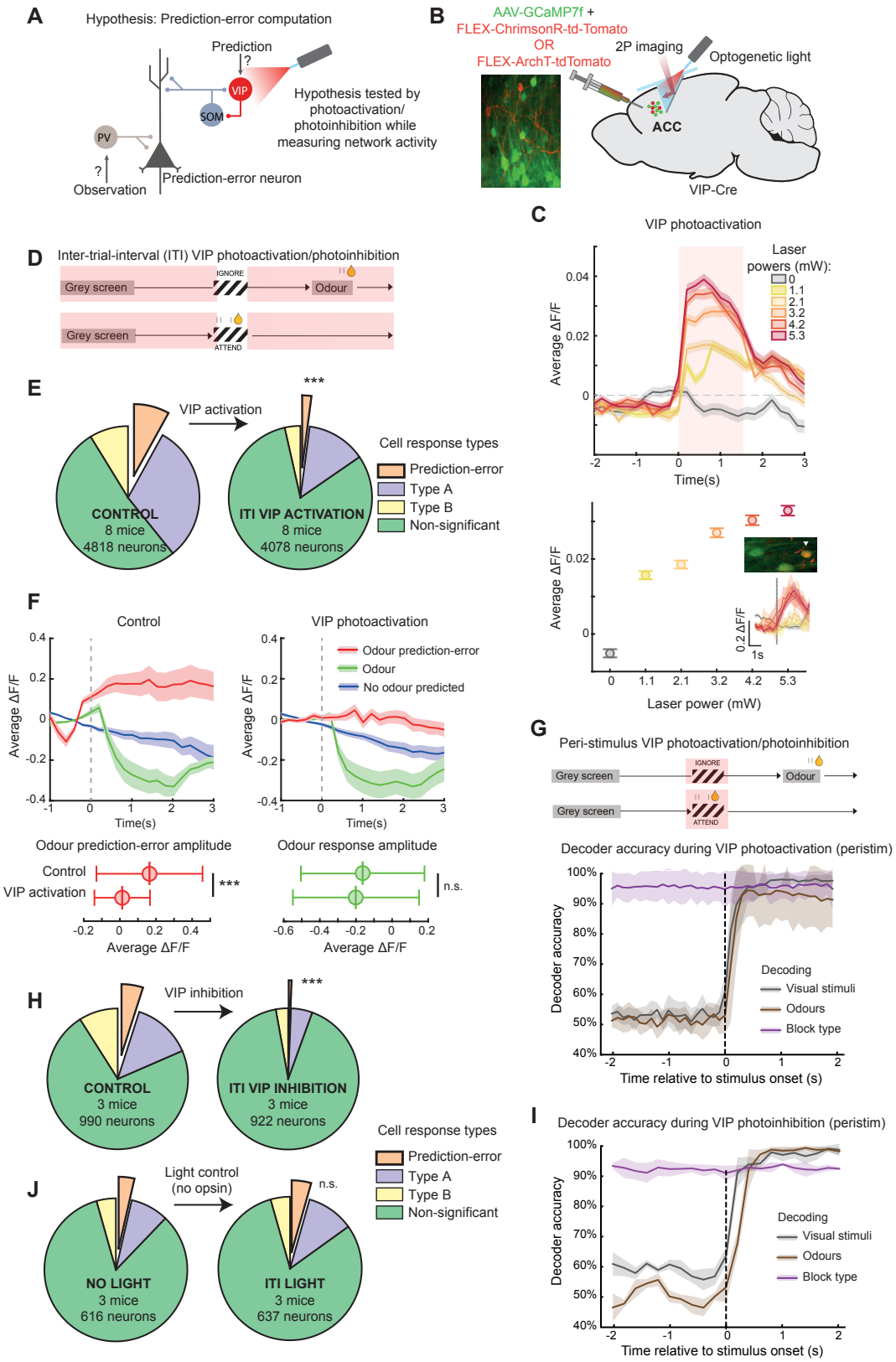



**Figure 7 | All-optical VIP interneuron bidirectional perturbations in ACC disrupt prediction-error signalling. A.** Hypothesis from theoretical work for a cortical circuit to compute prediction errors. The prediction-error neuron will respond (positive or negative responses) when predictions and

observations do not cancel each other out. Part of this hypothesis is tested by activating and inhibiting VIP interneurons during task-switching. **B**. Left, example site with neurons expressing GCaMP7f in green and VIP interneurons expressing the excitatory opsin ChrimsonR in red. Right, schematic of our all-optical approach: simultaneous 2-photon imaging and optogenetic activation or inhibition of VIP interneurons in the ACC. **C**. Top, peri-stimulus time histograms of mean activity of all non-VIP neurons (N = 2467 neurons, 8 mice) aligned to light onset, at 5 different light powers, for 1.5s. Shading and error bars indicate mean ± SEM here and below. Bottom, average activity of all neurons across light presentation window (0-1.5s). Inset: Image of an example VIP interneuron (arrowhead) and its response to increasing light powers, demonstrating photoactivation. **D**. Schematic of inter-trial-interval (ITI) activation epoch. **E**. Proportions of cells showing prediction-error responses and other cell classes as in Fig. 4H. Left, control and right, ITI VIP activation sessions from the same sites, performed on subsequent days. **F**. Top, average activity in each trial type of positively responding prediction-error cells identified from each session. Bottom, mean response (0 to 1.5s) of these cells to odour prediction error (left, red) and odour (right, green). Wilcoxon rank-sum tests comparing control vs ITI sessions, prediction-error responses $P = 1.1x10^{-4}$, odour responses P = 0.93. **G**. Top, schematic of peri-stimulus activation/inhibition 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. 4D. H. Proportions of cells showing prediction-error responses and other cell classes as in Fig. 4H. Left, control and right, ITI VIP inhibition sessions from the same sites, performed on subsequent days. **I**. Time course of decoding accuracy with a binary classifier using activity during peri-stimulus VIP inhibition sessions (N = 888 neurons, 3 mice). **J**. Proportions of cells showing prediction-error responses and other cell classes as in Fig. 4H. Left, no light and right, ITI light only control sessions from the same sites, performed on subsequent days.

## Discussion

In this study we demonstrate that mice can perform rapid and repeated block transitions in a cross-modal task switching paradigm. A theoretical RL model suggested that this behaviour was driven by a cognitive prediction-error signal, and in agreement with this prediction, we identified a neural signal which represented a cognitive prediction error in the ACC. Silencing and un-silencing the ACC in precise time windows demonstrated that this neural prediction-error signal was causally required for the mouse to update its behaviour between blocks. The amplitude of the prediction-error signal preceding a trial was related to whether or not a mouse would correctly switch its behaviour in the subsequent trial. Finally, VIP interneurons were identified as a key inhibitory cell class in the computation of cognitive prediction errors. Taken together, our results provide causal and mechanistic evidence for a longstanding idea, that the ACC computes cognitive prediction errors to guide flexible behaviour.

By training mice to perform repeated one-shot block transitions, we encapsulated a complete cognitive switch in a well-defined time window of a few seconds. Such rapid and complete cognitive transitions are akin to 'Aha! moments', where some information or insight allows an abrupt updating of one's internal model of the world [55]. These rapid cognitive transitions are difficult to study experimentally – one-shot transitions have primarily been observed in primates [34,35], and previous studies in rodents required tens to hundreds of trials to complete a behavioural transition [12,26–32,42]. Periods of intermediate accuracy which accompany these slower transitions prevent unambiguously relating a neural signal to its behavioural consequence. In our task, a well-defined prediction error leads to a demonstrable update in the animal's cognitive model in the next few seconds. Moreover, when studying the moment of cognitive transition, multiple block transitions are required in a single session for sufficient

608 statistical power. In addition, to be certain of an animal's internal model of the world requires
609 that the animal performs at high accuracy in the given task before the block transition,
610 typically resulting in longer and fewer blocks. Our task provided an unprecedented number
611 of repeated, highly rapid block transitions with mice switching between high accuracy
612 behaviours, which was crucial for conclusively assigning a behaviourally relevant role to the
613 neural prediction-error signal.

614 Cognitive prediction errors are distinct from prediction-error signals in other domains, such
615 as reward prediction errors in dopaminergic neurons, sensory prediction errors in sensory
616 cortex, or sensorimotor prediction errors in the cerebellum [2,56–59]. Cognitive prediction errors
617 enable flexible cognition, where one needs to maintain an abstract cognitive context or model
618 of the world in mind, which is updated when a prediction from the cognitive context or model
619 is violated. This is a distinct problem to solve compared to other prediction-error
620 computations and involves different brain regions [9]. Critically, the prediction-error responses
621 we found in this study did not resemble classical negative reward signals such as those found
622 in dopaminergic neurons. Instead, this signal resembles an outcome prediction error [52] which
623 relates to errors in predicting outcomes based on the currently believed rules of the task.

624 The computation of a prediction-error signal in our task requires comparing an internally
625 generated prediction signal with an external odour signal. The origin of the odour prediction
626 signal is currently unclear, and may reach the ACC from other prefrontal areas [11,60–62].
627 However, the ACC itself robustly represents the ongoing task rules as well as external stimuli,
628 and thus may compute the prediction error autonomously.

629 An important goal in understanding the neural circuit basis of cognition is to identify the
630 circuit which compares predictions with observations. While this circuit has been studied in
631 other contexts, such as for the dopaminergic reward-prediction system[57] or the visuomotor
632 mismatch system[56], it is poorly understood for cognitive rule-updating. In this study we took
633 advantage of a temporally well-defined cognitive prediction-error signal to take the first steps
634 in uncovering the circuit involved. We found VIP-driven disinhibition to be key, and expect
635 future studies to reveal further details about the role of other inhibitory cell classes in this
636 circuit.

637 The recurrently connected nature of cortical circuits may suggest that perturbation of any cell
638 class will invariably lead to a disruption of processing in that region. However, VIP modulation
639 can be orthogonal to cognitive modulations in other cortical regions[63], and importantly, here
640 we show that VIP perturbations in the ACC did not affect stimulus and context
641 representations, while disrupting prediction-error signals.

642 Other cortical association areas may play a role in task switching, such as orbitofrontal cortex
643 [12,27] and posterior parietal cortex [64], as well as subcortical areas including higher-order
644 thalamic nuclei [65]. Future work understanding the role of these and other regions, would need
645 to establish both the nature of the prediction-error signal and its requirement in behaviour.
646 Although we have ruled out a role for PL in our task, earlier studies have shown that PL circuits
647 are involved in flexible behaviour [45,66]. However, these studies investigated slower forms of
648 behavioural transitions, compared to our largely one-shot transitions, possibly accounting for
649 the difference with our results. Indeed, there are multiple types of surprise signals which may
650 be processed differentially by different brain regions [9,67].

Where is the prediction-error signal sent to in order to drive the behavioural changes? We have found that the striatum is an unlikely candidate, despite being a major projection target of the ACC, since the prediction-error signal is under-represented in ACC neurons projecting to the striatum. Other likely brain regions may include the locus coeruleus, which has been implicated in updating current strategies through norepinephrine signalling [17,18,68,69], or the dorsal raphe nucleus, which promotes behavioural flexibility through serotonin and glutamate signalling [70,71].

Disruption of cognitive flexibility in humans may lead to excessive transitioning between attentional states in ADHD, or excessive persistence in one or a few repetitive behaviours in ASD [72]. However, the neural circuit basis of these conditions is poorly understood. Our results provide a crucial insight regarding the role of the ACC in transitioning between, rather than maintaining, cognitive states. Thus, atypical ACC activity patterns may contribute to excessive or insufficient cognitive transitions in humans, and a more detailed understanding of how ACC circuits produce and transmit prediction errors may provide insights to better understand these conditions.

**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, MH and DM-J. All authors contributed to discussions and commented on the manuscript.

**Acknowledgments**

We thank John Duncan, Apoorva Bhandari, Dimitar Kostadinov and Sonja Hofer for discussions and comments on the manuscript, and Thomas Mrsic-Flogel, Petr Znamenskiy, Anil Seth, Jasper Poort and Juan Burrone for discussions of the results in this manuscript. We thank Marian 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. Keller, G. B. & Mrsic-Flogel, T. D. Predictive Processing: A Canonical Cortical Computation. *Neuron* **100**, 424–435 (2018).
3. Clark, A. Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behav Brain Sci* **36**, 181–204 (2013).
4. Bastos, A. M. *et al.* Canonical microcircuits for predictive coding. *Neuron* **76**, 695–711 (2012).
5. de Lange, F. P., Heilbron, M. & Kok, P. How Do Expectations Shape Perception? *Trends in Cognitive Sciences* **22**, 764–779 (2018).
6. Meirhaeghe, N., Sohn, H. & Jazayeri, M. A precise and adaptive neural mechanism for predictive temporal processing in the frontal cortex. *Neuron* **109**, 2995-3011.e5 (2021).
7. Sarafyazd, M. & Jazayeri, M. Hierarchical reasoning by neural circuits in the frontal cortex. *Science* **364**, eaav8911 (2019).
8. Shima, K. & Tanji, J. Role for cingulate motor area cells in voluntary movement selection based on reward. *Science* **282**, 1335–1338 (1998).
9. 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).
10. 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).
11. Bartolo, R. & Averbeck, B. B. Prefrontal Cortex Predicts State Switches during Reversal Learning. *Neuron* **106**, 1044-1054.e4 (2020).
12. Banerjee, A. *et al.* Value-guided remapping of sensory cortex by lateral orbitofrontal cortex. *Nature* **585**, 245–250 (2020).
13. 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.
14. Heilbronner, S. R. & Hayden, B. Y. Dorsal Anterior Cingulate Cortex: A Bottom-Up View. *Annu Rev Neurosci* **39**, 149–170 (2016).
15. Shenhav, A., Cohen, J. D. & Botvinick, M. M. Dorsal anterior cingulate cortex and the value of control. *Nat Neurosci* **19**, 1286–1291 (2016).
16. Kolling, N. *et al.* Value, search, persistence and model updating in anterior cingulate cortex. *Nat Neurosci* **19**, 1280–1285 (2016).
17. Tervo, D. G. R. *et al.* Behavioral variability through stochastic choice and its gating by anterior cingulate cortex. *Cell* **159**, 21–32 (2014).
18. Tervo, D. G. R. *et al.* The anterior cingulate cortex directs exploration of alternative strategies. *Neuron* **109**, 1876-1887.e6 (2021).
19. 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).
20. 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).
21. 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).
22. Kolling, N., Behrens, T. E. J., Mars, R. B. & Rushworth, M. F. S. Neural mechanisms of foraging. *Science* **336**, 95–98 (2012).
23. 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).
24. Quilodran, R., Rothé, M. & Procyk, E. Behavioral Shifts and Action Valuation in the Anterior Cingulate Cortex. *Neuron* **57**, 314–325 (2008).
25. Hyman, J. M., Holroyd, C. B. & Seamans, J. K. A Novel Neural Prediction Error Found in Anterior Cingulate Cortex Ensembles. *Neuron* **95**, 447-456.e3 (2017).

26. Jaramillo, S. & Zador, A. M. Mice and rats achieve similar levels of performance in an adaptive decision-making task. *Front. Syst. Neurosci.* **8**, (2014).
27. Liu, Y., Xin, Y. & Xu, N.-L. A cortical circuit mechanism for structural knowledge-based flexible sensorimotor decision-making. *Neuron* **109**, 2009-2024.e6 (2021).
28. Reinert, S., Hübener, M., Bonhoeffer, T. & Goltstein, P. M. Mouse prefrontal cortex represents learned rules for categorization. *Nature* (2021) doi:10.1038/s41586-021-03452-z.
29. Wang, T.-Y., Liu, J. & Yao, H. Control of adaptive action selection by secondary motor cortex during flexible visual categorization. *eLife* **9**, e54474 (2020).
30. Siniscalchi, M. J., Phoumthipphavong, V., Ali, F., Lozano, M. & Kwan, A. C. Fast and slow transitions in frontal ensemble activity during flexible sensorimotor behavior. *Nat. Neurosci.* **19**, 1234–1242 (2016).
31. Rodgers, C. C. & DeWeese, M. R. Neural Correlates of Task Switching in Prefrontal Cortex and Primary Auditory Cortex in a Novel Stimulus Selection Task for Rodents. *Neuron* **82**, 1157–1170 (2014).
32. Karlsson, M. P., Tervo, D. G. R. & Karpova, A. Y. Network resets in medial prefrontal cortex mark the onset of behavioral uncertainty. *Science* **338**, 135–139 (2012).
33. 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).
34. Harlow, H. F. The formation of learning sets. *Psychol Rev* **56**, 51–65 (1949).
35. Achterberg, J. *et al.* A One-Shot Shift from Explore to Exploit in Monkey Prefrontal Cortex. *J. Neurosci.* **42**, 276–287 (2022).
36. 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).
37. Pi, H.-J. *et al.* Cortical interneurons that specialize in disinhibitory control. *Nature* **503**, 521–524 (2013).
38. 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).
39. 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).
40. Letzkus, J. J., Wolff, S. B. E. & Lüthi, A. Disinhibition, a Circuit Mechanism for Associative Learning and Memory. *Neuron* **88**, 264–276 (2015).
41. Braga, A. & Schönwiesner, M. Neural Substrates and Models of Omission Responses and Predictive Processes. *Frontiers in Neural Circuits* **16**, (2022).
42. Akam, T. *et al.* The Anterior Cingulate Cortex Predicts Future States to Mediate Model-Based Action Selection. *Neuron* **109**, 149-163.e7 (2021).
43. Rushworth, M. F. S. & Behrens, T. E. J. Choice, uncertainty and value in prefrontal and cingulate cortex. *Nat Neurosci* **11**, 389–397 (2008).
44. 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).
45. 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).
46. Botvinick, M. M. Conflict monitoring and decision making: reconciling two perspectives on anterior cingulate function. *Cogn Affect Behav Neurosci* **7**, 356–366 (2007).
47. Passingham, R. E. *et al.* Attention to action. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences* **351**, 1473–1479 (1996).
48. 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).

49. 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).
50. Norman, K. J. *et al.* Frontal-Sensory Cortical Projections Become Dispensable for Attentional Performance Upon a Reduction of Task Demand in Mice. *Front Neurosci* **15**, 775256 (2022).
51. Zhang, S. *et al.* Organization of long-range inputs and outputs of frontal cortex for top-down control. *Nat. Neurosci.* **19**, 1733–1742 (2016).
52. Alexander, W. H. & Brown, J. W. Medial prefrontal cortex as an action-outcome predictor. *Nat Neurosci* **14**, 1338–1344 (2011).
53. Kuchibhotla, K. V. *et al.* Parallel processing by cortical inhibition enables context-dependent behavior. *Nat. Neurosci.* **20**, 62–71 (2017).
54. Attinger, A., Wang, B. & Keller, G. B. Visuomotor Coupling Shapes the Functional Development of Mouse Visual Cortex. *Cell* **169**, 1291-1302.e14 (2017).
55. Kounios, J. & Beeman, M. The Aha! Moment: The Cognitive Neuroscience of Insight. *Curr Dir Psychol Sci* **18**, 210–216 (2009).
56. Leinweber, M., Ward, D. R., Sobczak, J. M., Attinger, A. & Keller, G. B. A Sensorimotor Circuit in Mouse Cortex for Visual Flow Predictions. *Neuron* **95**, 1420-1432.e5 (2017).
57. Watabe-Uchida, M., Eshel, N. & Uchida, N. Neural circuitry of reward prediction error. *Annu Rev Neurosci* **40**, 373–394 (2017).
58. Rabinovich, R. J., Kato, D. D. & Bruno, R. M. Learning enhances encoding of time and temporal surprise in mouse primary sensory cortex. *Nat Commun* **13**, 5504 (2022).
59. Sokolov, A. A., Miall, R. C. & Ivry, R. B. The Cerebellum: Adaptive Prediction for Movement and Cognition. *Trends in Cognitive Sciences* **21**, 313–332 (2017).
60. Wilson, R. C., Takahashi, Y. K., Schoenbaum, G. & Niv, Y. Orbitofrontal Cortex as a Cognitive Map of Task Space. *Neuron* **81**, 267–279 (2014).
61. 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).
62. 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.
63. Myers-Joseph, D., Wilmes, K. A., Fernandez-Otero, M., Clopath, C. & Khan, A. G. Disinhibition by VIP interneurons is orthogonal to cross-modal attentional modulation in primary visual cortex. *Neuron* **112**, 628-645.e7 (2024).
64. Fox, M. T., Barense, M. D. & Baxter, M. G. Perceptual attentional set-shifting is impaired in rats with neurotoxic lesions of posterior parietal cortex. *J Neurosci* **23**, 676–681 (2003).
65. Marton, T., Seifikar, H., Luongo, F. J., Lee, A. T. & Sohal, V. S. Roles of prefrontal cortex and mediodorsal thalamus in task engagement and behavioral flexibility. *J. Neurosci.* (2018) doi:10.1523/JNEUROSCI.1728-17.2018.
66. Cho, K. K. A., Shi, J., Phensy, A. J., Turner, M. L. & Sohal, V. S. Long-range inhibition synchronizes and updates prefrontal task activity. *Nature* **617**, 548–554 (2023).
67. Modirshanechi, A., Brea, J. & Gerstner, W. A taxonomy of surprise definitions. *Journal of Mathematical Psychology* **110**, 102712 (2022).
68. 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).
69. Jordan, R. & Keller, G. B. The locus coeruleus broadcasts prediction errors across the cortex to promote sensorimotor plasticity. *Elife* **12**, RP85111 (2023).
70. 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).
71. Ahmadlou, M. *et al.* A subcortical switchboard for exploratory, exploitatory, and disengaged states. 2023.12.20.572654 Preprint at https://doi.org/10.1101/2023.12.20.572654 (2023).
72. 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.

73. Couto, J. *et al.* Chronic, cortex-wide imaging of specific cell populations during behavior. *Nat Protoc* **16**, 3241–3263 (2021).
74. Poort, J. *et al.* Learning Enhances Sensory and Multiple Non-sensory Representations in Primary Visual Cortex. *Neuron* **86**, 1478–1490 (2015).
75. Poort, J. *et al.* Learning and attention increase visual response selectivity through distinct mechanisms. *Neuron* **110**, 686-697.e6 (2022).
76. Pachitariu, M. *et al.* Suite2p: beyond 10,000 neurons with standard two-photon microscopy. 061507 Preprint at https://doi.org/10.1101/061507 (2017).
77. Mathis, A. *et al.* DeepLabCut: markerless pose estimation of user-defined body parts with deep learning. *Nat Neurosci* **21**, 1281–1289 (2018).
78. Li, N. *et al.* Spatiotemporal constraints on optogenetic inactivation in cortical circuits. *Elife* **8**, (2019).

848 **METHODS**

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

851 **Surgical Procedures**

852 For all surgeries, mice were anaesthetised using isoflurane, at 4% concentration for induction and at
853 0.5-1% for maintenance. Additional drugs were used to provide analgesia (Metacam 5mg/kg), anti-
854 inflammatory effects (dexamethasone 3.8mg/kg), and to reduce mucus secretions (Atropine
855 0.08mg/kg). Eye-cream (Maxitrol) was applied to the eyes to prevent drying and body temperature
856 was maintained at 37$^{o}$C using a heating mat and rectal temperature probe (Harvard Apparatus).
857 Injections of antibiotic (Betamox 120mg/kg) and analgesia (methadone hydrochloride 10mg/kg) were
858 given before the withdrawal of anaesthesia, and further analgesia was given daily for 1-2 days during
859 recovery of the animal.

860 For the two-photon imaging experiments, 10 VIP-Cre mice (C57BL/6-VIP$^{tm1(cre)zjh}$, The Jackson
861 Laboratory, strain # 010908, P42-49, 6 males and 4 females) were used. An additional 4 mice of the
862 same genotype (1 male, 3 females) were used together with the imaging mice for behavioural analysis.
863 A circular piece of scalp was removed and the underlying skull was cleaned. Small holes were drilled
864 in the skull above injection sites, located using stereotaxic co-ordinates. Injections of a mixture of
865 viruses expressing GCaMP7f (pGP-AAV9-syn-jGCaMP7f-WPRE, Addgene) and Cre-dependent
866 ChrimsonR (pAAV5-syn-FLEX-rc[ChrimsonR-tdTomato], Addgene) were made in the anterior cingulate
867 cortex (ACC, +0.9mm AP, -1.3 rising to -0.8mm DV) of either the left or right hemisphere (+/- 0.55mm
868 ML), using glass pipettes and a pressure micro-injection system (Picospritzer III, Parker). An injection
869 of Cholera Toxin Subunit B (recombinant) conjugated to Alexa Fluor 647 (ThermoFisher) was made in
870 the striatum of the same hemisphere (+1.2mm AP, +/- 1.5mm ML, -2.5 rising to -2.0mm DV) to
871 retrogradely label cells projecting to the striatum. 8 out of the 10 mice had good enough quality z-
872 stacks in the ACC to visualise CTB-Alexa647 and were used to identify striatal projecting ACC neurons.
873 8 out of the 10 mice exhibited sufficiently consistent behaviour to study the effects of VIP
874 photoactivation. For the VIP silencing experiments (Figure 7H) 3 VIP-Cre mice were injected with a
875 mixture of viruses expressing GcaMP7f (pGP-AAV9-syn-jGCaMP7f-WPRE, Addgene) and Cre-
876 dependent ArchT (pAAV-FLEX-ArchT-tdTomato (AAV5), Addgene). For the light-only control
877 experiments (Figure 7J), an additional 3 VIP-Cre mice were injected with a mixture of viruses
878 expressing GCaMP7f and Cre-dependent tdTomato (pAAV-FLEX-tdTomato (AAV1), Addgene).

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

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

893 Pacer Technology) were thinly applied onto the cement to increase the transparency. After recovery
894 from surgery for at least 5 days mice began habituation and behavioural training.

895 For the optogenetic silencing experiments, 8 transgenic mice (P42-49, 4 males and 4 females)
896 expressing Channelrhodopsin-2 in parvalbumin-expressing interneurons were used, obtained by
897 crossing FLEX ChR2 mice (Gt(ROSA)26Sor$^{tm32(CAG-COP4*H134R/EYFP)Hze}$) and PV-Cre mice (Pvalb$^{tm1(cre)Arbr}$, The
898 Jackson Laboratory, strain # 204109 and # 017320 respectively). Two small holes were drilled above
899 the ACC and prelimbic cortex (PL) of each hemisphere. Dual-core cannulae with bilateral optical fibres
900 (Thorlabs), each with a diameter of 200µm and 0.39NA, cut to a length of < 3mm, were implanted in
901 the ACC (+0.9mm AP, -1.2mm DV, +/-0.35mm ML) and PL (+2.6mm AP,-1.25mm DV, +/- 0.35mm ML),
902 and the stainless steel ferrules were bonded to the skull using dental cement (C&B Superbond), along
903 with a custom machined head-plate. PL implants were inserted at a 25° angle (relative to vertical)
904 through holes drilled 0.8mm anterior to PL. We performed light-only control experiments by
905 comparing 3 PV-Cre mice (1 male and 2 females) with 3 PV-Cre crossed with FLEX-ChR2 mice (1 male
906 and 2 females), which were implanted with optical fibre cannulae above ACC only. After recovery from
907 surgery for at least 5 days mice began habituation and behavioural training.

908 **Imaging**

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

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

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

928 Widefield imaging was performed on a custom built inverted tandem lens macroscope (Cosys), with
929 two photographic lenses (AF-S NIKKOR 85mm f/1.8G lens and AF NIKKOR 50mm f/1.4D Lens). The
930 brain was illuminated with interleaved collimated blue (470nm, Thorlabs M470L4) and violet light
931 (405nm, Thorlabs M405L4) at an irradiance of ~0.03mW/mm$^2$. Images were recorded with a CMOS
932 camera (Point Grey Research Grasshopper3) at frame rate of 54Hz. LEDs and camera frame acquisition
933 were triggered using a digital microprocessor (Teensy 3.2).

934 Widefield data was pre-processed using the methods described in [73]. The widefield video underwent
935 motion correction and the brain images were aligned within and across mice by manual rigid
936 alignment to a number of anatomical landmarks. The video data was compressed and denoised by

937 performing SVD on the matrix of pixels × time and retaining the top 500 components. The ΔF/F was
938 computed for each pixel by taking the difference between F and F0, and dividing by F0, where F0 was
939 the mean value across the entire session. Traces were filtered with a 0.0033 Hz high-pass second order
940 Butterworth filter, and an additional 7Hz lowpass filter was applied to the violet illumination trace. To
941 correct for haemodynamic artefacts, a scaled version of the violet illumination trace was subtracted
942 from the blue illumination trace for each pixel. This scaling factor was found by regressing the violet
943 trace onto the blue trace. To account for overt movement-related brain activity, we fit a ridge
944 regression model to the data, predicting brain activity from a number of movement regressors. These
945 included a binarized lick trace, with lags up to 500ms as well as instantaneous running speed and
946 average face motion energy. Running speed and face motion energy were divided by twice their
947 standard deviation to ensure all regressors had approximately the same scale and were penalised
948 equally by ridge regularisation. Ridge penalties were selected using fivefold cross-validation from 36
949 values spaced logarithmically between $10^{-2}$ and $10^{5}$ selecting the ridge penalty which resulted in the
950 lowest cross validated mean squared error. Penalties were selected independently for each pixel. We
951 then subtracted this predicted activity and all subsequent analysis was performed on the model
952 residuals.

953 **Behavioural training**

954 The behaviour apparatus and training was similar to previous studies [74,75]. Mice were trained on a
955 visual discrimination task for up to two weeks, until discrimination performance reached threshold,
956 before training them on the switching task (see below). Mice had free access to water but were food
957 restricted to maintain at least 85% of their free-feeding body weight (typically 85-90%, 2-3g of
958 standard food pellets per animal per day). A reward delivery spout was positioned near the snout of
959 the mouse, and licks to this spout were detected using a piezo disc sensor and custom electronics. The
960 reward was a 10% solution of soy milk powder (SMA Wysoy) delivered by opening a pinch valve
961 (NResearch) controlled through custom electronics. The mouse's running speed on the cylinder was
962 measured using an incremental rotary encoder (Kübler). Two luminance-corrected monitors
963 (luminance meter LS-100, Konica Minolta) positioned at 45° angles and 25cm distance to the mouse
964 delivered visual stimuli.

965 Animals were habituated to handling and gentle restraint over two to three days, before they were
966 head-fixed and trained to run on a polystyrene cylinder (20cm diameter) for one to four days. This
967 period was also used to find suitable imaging sites. After the habituation phase, mice performed one
968 behaviour session in which the movement of the gratings was linked to the mouse's movement on the
969 wheel. Subsequently, mice were trained to self-initiate trials by sustained running on the wheel for at
970 least 2.8s and an added random duration drawn from an exponential distribution with mean 0.4s (trial
971 structure and all timings shown in Supplementary Figure 1A). At this point one of two drifting
972 sinusoidal visual gratings were randomly presented, drifting in the opposite direction to the direction
973 of running, with a fixed spatial and temporal frequency of 0.1 cycles per degree and 2Hz respectively.
974 The rewarded and unrewarded gratings were oriented +/-20° relative to vertical, symmetrically on
975 both screens. When the rewarded grating was displayed the mouse could trigger the delivery of a
976 reward, a drop of soya milk, by licking the spout during the 'reward period', lasting from 1.0s after the
977 appearance of the grating to its disappearance, maximum 0.8 s into the reward period. This was
978 recorded as a 'hit'. In some sessions, the duration of the reward period started at 1.5s and lasted up
979 to 1.53 seconds, no difference in behaviour was observed with this minor change in timings. The visual
980 stimulus stayed on for an additional 0.8s after reward onset while the mouse consumed the reward.
981 If the mouse did not lick during this period, the trial was recorded as a 'miss', and a drop of soy milk
982 was delivered shortly before the disappearance of the grating. When the unrewarded grating was

983 presented, a single lick or more at any time until the stimulus disappearance was recorded as a 'false
984 alarm', triggering a time-out period of 4s in which the unrewarded grating remained on screen, and
985 any further licks restarted the time-out. During early training the probability of unrewarded trials was
986 occasionally increased transiently up to 0.8 to discourage erroneous licking. All mice learned the visual
987 discrimination task in 5-10 days, with post-learning defined as three consecutive days of discrimination
988 with a behavioural d-prime score of 2 or above. Behavioural d-prime was calculated as: $bd' = \Phi^{-1}(H) -$
989 $\Phi^{-1}(F)$, where $\Phi^{-1}$ is the normal inverse cumulative distribution function, H is the rate of hit trials, and
990 F is the rate of false alarm trials.

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

998 Once animals were discriminating the odours accurately (typically after 30-40 trials), they were trained
999 to switch between blocks of the olfactory and visual discrimination task. Mice typically learned to
1000 switch successfully in 1-3 days. In the olfactory blocks, 70% of odour stimuli were preceded with one
1001 of the two same visual gratings featured in the visual discrimination task (fixed duration of 1.8s, with
1002 an identical onset delay distribution as in the visual block). In this instance neither grating was
1003 rewarded or punished, and mice learned to ignore these irrelevant gratings while accurately
1004 discriminating the odours, which were presented after the irrelevant visual grating (delay between
1005 visual grating offset and odour onset 1.8s (in some two-photon imaging sessions this delay was 1.8s
1006 plus an added random duration drawn from an exponential distribution with mean 0.2s). In initial
1007 switching training sessions, a reward was delivered at the end of a rewarded grating in a visual block
1008 if the mouse had failed to lick, giving a clear indication that the grating was now relevant. By the end
1009 of early training, and for all data in this paper except Supplementary Figure 2F, this feature was
1010 removed, requiring mice to switch between blocks through noticing unexpected stimuli alone. Block
1011 switches occurred automatically when a mouse had demonstrated a > 80% discrimination
1012 performance to the relevant stimuli (visual gratings in visual block, odours in odour block) over the
1013 last 30 trials of a block. Additionally in odour blocks mice were required to have successfully ignored
1014 all irrelevant visual gratings over the previous 10 trials before a block transition was triggered. Blocks
1015 typically contained 30 to 40 trials. Mice were deemed to have learned the switching task when they
1016 could complete sessions at these parameters with at least 3 repeats of each block type.

1017 In order to compare the speed of behavioural switching between blocks, we applied a transition period
1018 immediately after a block transition where visual stimuli were selected according to the following
1019 rules. For odour to visual block transitions, in the first trials of a visual block only the rewarded grating
1020 (visual stim 1) was shown. When a mouse responded correctly by licking to these grating stimuli on
1021 three consecutive trials this transition period ended and the block continued with the normal 50%
1022 probabilities of visual grating identities. In the visual to odour block transitions, we applied two
1023 variations of the transition period. In the first variation, the first irrelevant visual stimulus was the
1024 otherwise unrewarded visual stimulus 2, ensuring that that the block transition was indicated by the
1025 unexpected appearance of an odour, rather than a reward prediction error. The subsequent irrelevant
1026 stimuli were visual stimulus 1, and the transition period ended when a mouse responded correctly by
1027 not licking to these irrelevant visual stimuli on three consecutive trials. Odour stimuli selection itself
1028 was kept random. In the second variation, even the first irrelevant visual stimulus was visual stimulus

1, and the subsequent rules were the same. We confirmed that mice switched equally fast in both variants of the task (Wilcoxon rank sum test, P > 0.05). These transition periods were used in all behaviour sessions except muscimol silencing experiments (Supplementary Figure 2C) and light-only controls (Supplementary Figure 2B), 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. Code for, and experimental data used to train the models is available at https://github.com/adityagilra/BeliefStateRL. 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 the 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 the 2 visual stimuli were shown. In step 2, a needless lick was punished (-1), and one of the 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. Overall, a correct response for a trial in either block is defined as a lick for cue 1 or a no lick for cue 2 in the final reward step, while requiring no-lick in all other time steps. Block switches occurred using the same transition rules as in the experiment, described above.

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 $\epsilon$-greedy policy. i.e. a random action was taken with probability $\epsilon$, 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 $\alpha$

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

Cues $s', s$ and actions $a', a$ correspond to 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 $\epsilon$-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 $\delta$ multiplied by a learning rate $\alpha$.

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 $\xi$ was a Gaussian random variable with mean 0 and variance 1, and $\beta$ was a noise factor parameter. The agent's belief was updated as:

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

where $\chi'$ was multiplied by a belief switching rate $\zeta$, 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 for the next trial. After training, updating of Q-value tables was turned off, and only belief updates were carried out.

Model fitting: Each model was simulated with learning via SARSA for $N/2$ time steps. After this training, we obtained simulated behaviour data for a further $N/2$ time steps, for fitting to experimental data of trained mice. For Basic RL model fitting, exploration and updation of Q-value tables via SARSA parametrized by $\epsilon$ and $\alpha$ were kept the same as during training. For Belief-state RL model fitting, Q-values were no longer updated i.e. $\alpha = 0$ after training, only the belief state was updated, and exploration was kept on. We confirmed that keeping Q-value updating on after learning, at $\alpha = 0.1$ as during training, did not have a noticeable effect on the results.

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. RMSE mean for each fold i.e. each training and test split was calculated across 5 seeds (2 with the first variation + 3 seeds with the second variation in a similar ratio as the number of experimental sessions on the two variations of the block transitions described above). To select between the models, the RMSE mean $\pm$ SD across these 5 folds, computed on the above RMSE mean across seeds, were reported and compared as below.

For the default basic RL model, we fit 2 parameters: exploration rate $\epsilon$ and the learning rate $\alpha$. 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 default belief-state RL model, we fit 4 parameters: belief switching rate $\zeta$, the noise factor $\beta$ for the prediction-error signal, exploration rate $\epsilon$, and a different exploration rate $\epsilon'$ to account for enhanced licking to visual cue 2 in the olfactory block. The learning rate $\alpha$ was fixed at 0.1 during training, but was set to zero for the fitted data, as the belief switching rate played a much stronger role in rapid switching between blocks. The RMSE mean $\pm$ SD on training

and test splits were 0.157438 $\pm$ 0.004906 and 0.193520 $\pm$ 0.010869 for the default basic RL agent with 2 parameters, and 0.087012 $\pm$ 0.005060 and 0.137714 $\pm$ 0.014658 for the default belief-state RL with 4 parameters, leading us to choose the Belief-state RL model over the Basic one.

We also fitted a Basic RL model with 3 parameters: exploration rate $\epsilon$, learning rate $\alpha$, and independent exploration rate $\epsilon'$ on receiving visual cue 2, which yielded RMSE mean $\pm$ SD on 5-fold training and test splits as 0.161927 $\pm$ 0.006923 and 0.197955 $\pm$ 0.011329. Further, we fitted a Belief-state RL model with 3 parameters: belief switching rate $\zeta$, noise factor $\beta$, and a common exploration rate $\epsilon$, which yielded RMSEs as 0.095403 $\pm$ 0.006537 and 0.143631 $\pm$ 0.015731. This shows that $\epsilon'$, the enhanced exploration rate to visual cue 2, does not play a major role in selecting between Basic vs Belief-state RL models.

Simulated $p(lick|cue)$ using the best fit parameters (on the full dataset) for each model for one seed are shown in Figure 1G and Supplementary Figure 3C middle and bottom. The prediction-error signal for one-shot versus slower switches made by the agent, shown in Figure 6C, was computed as the sum of the absolute values of the two components of the prediction-error signal $\chi$ 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 (Supplementary Figs. 2d and 3f), using two parameters – factors on the prediction-error signal for odour and visual trials. These factors signified how much the prediction-error signal reduced on silencing the ACC.

Fitted parameters are shown in Table 1. Since these fits were not for model selection, all of the data was fit, minimizing mean RMSE across 5 seeds (both variations of the task included as described above).

| 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) | $\epsilon, \alpha$ = 0.236, 0.9 | 0.158488 ± 0.018537 |
| Belief-state RL on primary dataset | *N* = 500,000 $\alpha = 0.1$ during training, $= 0$ during testing | $\zeta, \beta, \epsilon, \epsilon' =$ 0.6858, 1.997, 0.049, 0.280 | 0.080526 ± 0.004806 |
| Belief-state RL on ACC silencing dataset with ACC not silenced | N=500,000 $\alpha = 0.1$ during training, $= 0$ during testing | $\zeta, \beta, \epsilon, \epsilon' =$ 0.7698, 1.987, 0.1072, 0.4516 | 0.053635 ± 0.001487 |

| | | | |
|---|---|---|---|
| Belief-state RL on ACC silencing dataset with ACC silenced | N=500,000<br>$\alpha = 0.1$ during training, = 0 during testing<br>$\zeta, \beta, \epsilon, \epsilon' =$ 0.7698, 1.987, 0.1072, 0.4516 | ACC inhibition factors on prediction error for visual and odour trials = 0.2143, 0.5016 | 0.0742734 ± 0.000907 |

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 neuron in each session using Suite2p [76]. Each site yielded between 129 and 925 neurons, median = 499 neurons. Raw fluorescence time series F(t) were obtained for each neuron by averaging across pixels within each ROI. Baseline fluorescence F0(t) was computed by smoothing F(t) (causal moving average of 0.75s) and determining for each time point the minimum value in the preceding 60 s time window. The change in fluorescence relative to baseline, ΔF/F, was computed by taking the difference between F and F0, and dividing by F0.

To identify prediction-error neurons, we selected neurons which responded significantly differently to the odour prediction-error event when compared to both the expected arrival of the odour and when no odour was expected (Figure 2B). We defined three epochs each lasting 1.5s and measured the average neural activity in these epochs: 1) Odour prediction-error trials, starting 2.0s after the offset of the first visual stimulus following a switch from an odour block to a visual block, provided the mouse did not already lick to the preceding visual stimulus (2.0s was the average delay in the imaging sessions from the visual stimulus offset to the odour stimuli onset). 2) Stable odour block trials from the end of the preceding odour block (when an odour is expected and received) aligned to the onset of the odour stimuli, following a correctly ignored visual grating. 3) No odour expected trials, when no odour is expected following a visual stimulus, 2.0s following the offset of an unrewarded visual stimulus. Trials were taken from subsequent visual block up to 10 trials before the end of the block. In epochs 2 and 3, we averaged up to 7 trials of each condition for each block transition (median 7 trials). We compared the neural activity in different epochs with a Wilcoxon rank-sum test with the the number of samples equal to the number of block transitions. Prediction-error neurons were defined as neurons with significantly different activity in odour prediction-error trials, when compared to both of the other conditions. We repeated the analysis without averaging multiple trials in epochs 2 and 3 and still obtained a significantly larger fraction of prediction-error neurons in the ACC compared to V1 (Chi-squared test of proportions $P < 0.0001$). Similar criteria were used for the visual to odour block transitions. Positively and negatively responsive prediction-error neurons were those in which the response to the odour prediction error 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-error condition (99 neurons), and second, the reverse of this (163 neurons). To test whether activity in prediction-error neurons could predict subsequent switching, average activity between 0-1.5s in odour prediction-error 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.

1176 Decoding analysis (Figure 4D, 7G, I) was performed by training a binary logistic regression classifier to
1177 decode the stimulus or block identity from the vector of ΔF/F values at each frame. Decoding
1178 performance was assessed using stratified 5-fold cross validation and taking the mean accuracy across
1179 the 5 test sets. Stimuli classes were evenly balanced by randomly subsampling the larger class. We
1180 applied an L2 regularisation penalty to reduce overfitting.

1181 For identification of striatal projecting neurons, a brief dual channel recording of the imaging planes
1182 was acquired before each imaging session at an excitation wavelength of 830nm. Following
1183 segmentation, imaged neurons co-expressing Alexa-647 were identified using this recording, and
1184 confirmed using a detailed anesthetised dual channel z-stack taken at the end of all imaging sessions.
1185 To calculate confidence intervals for the percentage of prediction-error neurons in the non-striatal
1186 projecting ROIs (Figure 4M) a percentile bootstrap method was used, resampling with replacement a
1187 number of ROIs equivalent to the size of the striatal projecting ROI population 10,000 times. The
1188 proportion of prediction-error neurons was then calculated and the $0.5^{th}$ and $99.5^{th}$ percentiles of this
1189 distribution of proportions was calculated to obtain the 99% confidence interval.

1190 To assess the proportion of neural activity which was attributable to overt behaviour recorded during
1191 our task (Supplementary Figure 5C, D, E) a linear model was fit using ridge regression to predict neural
1192 activity. The model was constructed by combining multiple sets of variables into a design matrix, to
1193 capture signal modulation by the following different task or behavioural events: 2 visual stimuli, 2
1194 odour stimuli, reward delivery, licks, running speed, block type, and an interaction term for visual
1195 stimuli and block type. Each stimulus/event variable was structured to capture a time-varying event
1196 kernel. Variables therefore consisted of a vector of the relevant stimulus/event, and copies of this
1197 vector, each shifted in time by one frame for specific durations. For sensory stimuli, the time-shifted
1198 copies ranged up to 2s after the original. For motor events (running and licking) the time-shifted copies
1199 spanned the frames from 0.5s before until 2s after the original. The model was fit with 5-fold cross
1200 validation and the coefficient of determination ($R^2$) was calculated based on the predictions of the
1201 model on held out data not used during training. We then assessed the predictive power of the
1202 behavioural model variables by comparing the $R^2$ value for the full model to a model without the
1203 running and licking predictors.

1204

1205 **Optogenetic activation and inhibition of VIP interneurons during ACC imaging**

1206 To perturb VIP neurons expressing Chrimson, ArchT, or tdTomato concurrently with 2-photon imaging,
1207 a 639nm laser (OBIS, Coherent) was used to deliver light via a 200µm diameter, 0.39 NA optic fibre
1208 (Thorlabs) positioned around 3mm from the posterior edge of the microprism, at a 30° angle relative
1209 to the coverslip. The laser and stimulus monitors were blanked during the linear phase of the resonant
1210 scanner to allow quasi-simultaneous two-photon imaging and optogenetic activation. The effective
1211 maximum output power from the optic fibre was 5.3mW. During an optogenetic calibration session in
1212 the dark, 5 laser powers (20%, 40%, 60%, 80%, and 100% of maximum) were pseudorandomly applied
1213 to the coverslip for 1.5s, with a 5 second interval between each. During switching sessions, two laser
1214 epochs were used, with 5.3mW power only. In the peri-stimulus epoch, the laser began 0.1s before
1215 the visual stimulus and continued until the end of the visual stimulus. In the inter-trial-interval epoch,
1216 the laser began at the offset of each visual stimulus and continued until the onset of the next visual
1217 stimulus.

1218 **Pupil tracking**

1219 Eye recordings were acquired using a monochrome USB2.0 video camera (The Imaging Source) with a
1220 50mm 2/3" format 5-megapixel lens (Computar), set to acquire at 320x240 (Y800) resolution and 30
1221 frames per second. Frames were triggered using an Arduino Uno microcontroller board (Arduino) to
1222 ensure a constant acquisition rate. Pupil data were extracted using DeepLabCut [77] for 2D marker
1223 tracking, with markers set to track vertical and horizontal boundaries of the eye and pupil throughout
1224 the recording. Frames that coincided with blinks were removed based on changed in vertical size of
1225 the eye, and pupil width was calculated from the remaining frames.

1226 **Optogenetic silencing of ACC activity**

1227 Once mice had learned the full switching task, optogenetic silencing of ACC neurons was performed
1228 by connecting the optic fibre cannulae to a blue LED (470nm, Thorlabs), and delivering light while the
1229 mouse performed the task. Before implantation, each optic fibre was confirmed to have an effective
1230 power output of >1mW after cutting. Light was delivered either throughout the session (pulsed at
1231 40Hz [78]), 0.5s before each visual stimulus continuing to the end of the stimulus ('peri-stimulus'), or
1232 from the end of each visual stimulus until the beginning of the next visual stimulus ('inter-trial-
1233 interval', ITI). These three epochs were used to silence both ACC and PL on different days, creating 6
1234 silencing conditions. These conditions were pseudorandomly chosen across consecutive days, with
1235 the order different between mice but the time between repeated conditions kept constant, and with
1236 control no-light sessions interspersed every third session. For one-trial un-silencing, the light was
1237 continuously pulsed throughout the session as above, but this was paused at the end of the last trial
1238 in the odour block, and resumed at the start of the second trial of the visual block. In the peri-stimulus
1239 and ITI epochs, light power was ramped down for the final 200ms of each pulse. For ACC silencing
1240 during learning, the silencing group included all 8 optogenetic mice, and the non-silenced controls
1241 were the 10 imaging mice. Light-only controls were performed in PV-Cre mice not expressing
1242 channelrhodopsin (N = 3 mice). No differences were found in these mice in the number of trials to
1243 transition between blocks with and without the light stimulation ($P > 0.05$, Wilcoxon rank sum test).

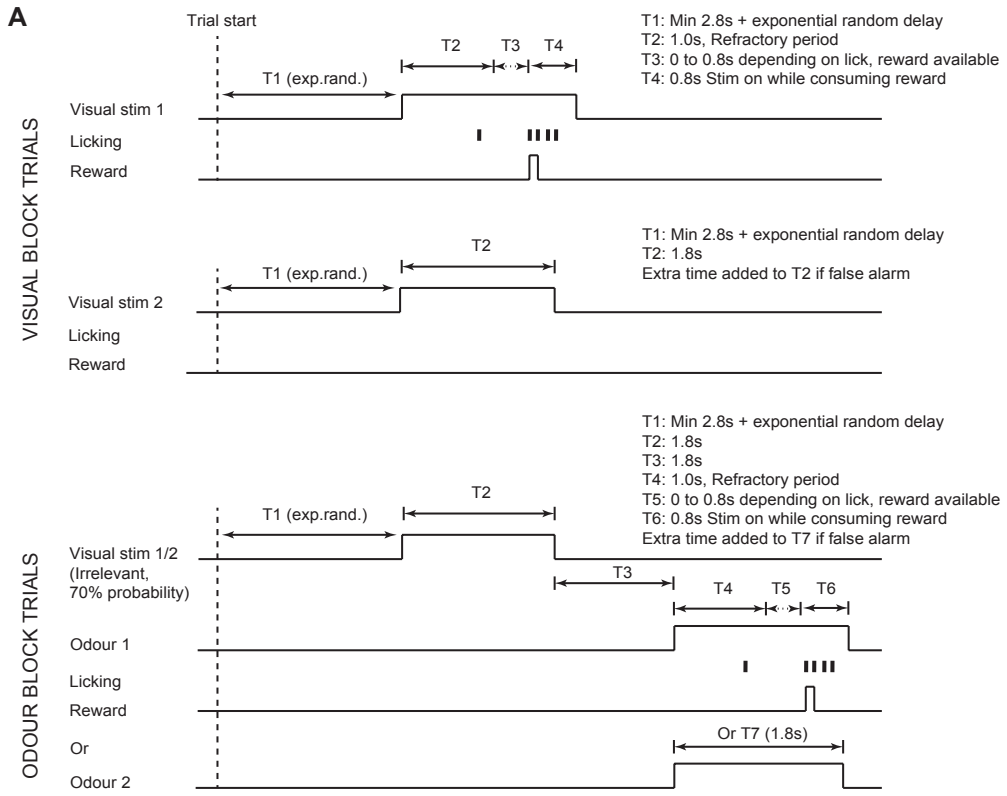
1244 **Pharmacological silencing of ACC activity**

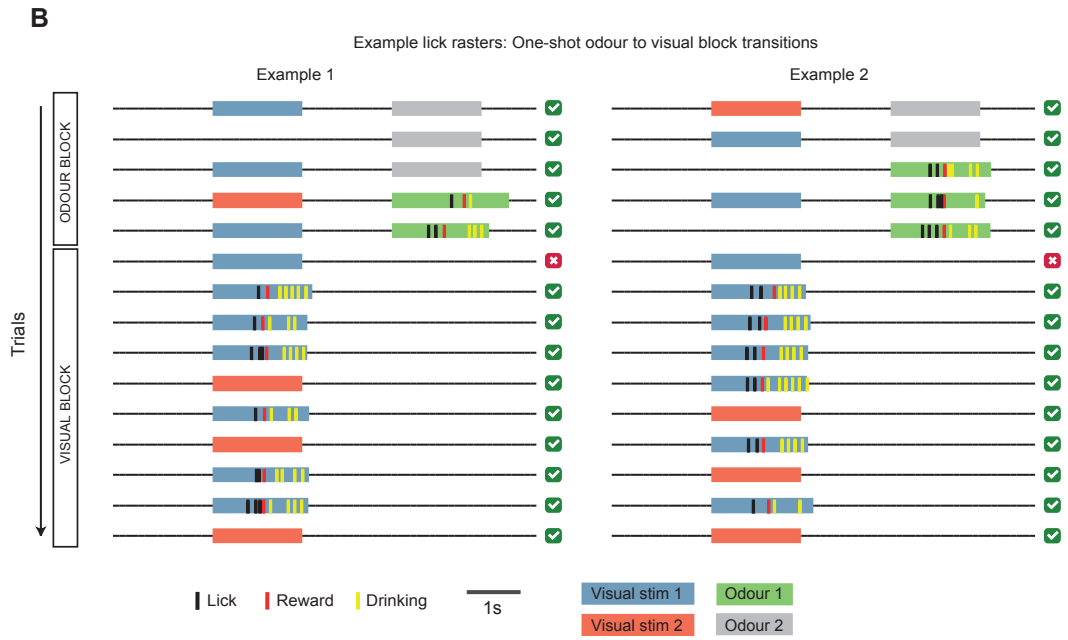
1245 For the muscimol silencing experiments, 4 male wildtype C57Bl/6j mice were implanted with bilateral
1246 infusion cannulae (-1.2mm DV, +0.7mm AP, ± 0.5mm ML, bilaterally). After training in the switching
1247 task, mice were infused bilaterally with 300nl muscimol (Sigma, 1µg/ µl) or saline at a rate of
1248 0.25µl/min using a 1µl syringe (Hamilton) and syringe pump (World Precision Instruments SP100IZ) 30
1249 minutes before the start of a switching session. We waited 5 minutes after the syringe pump had
1250 finished to allow full infusion of the drug before disconnecting the cannulae.

1251 **Silencing data analysis**

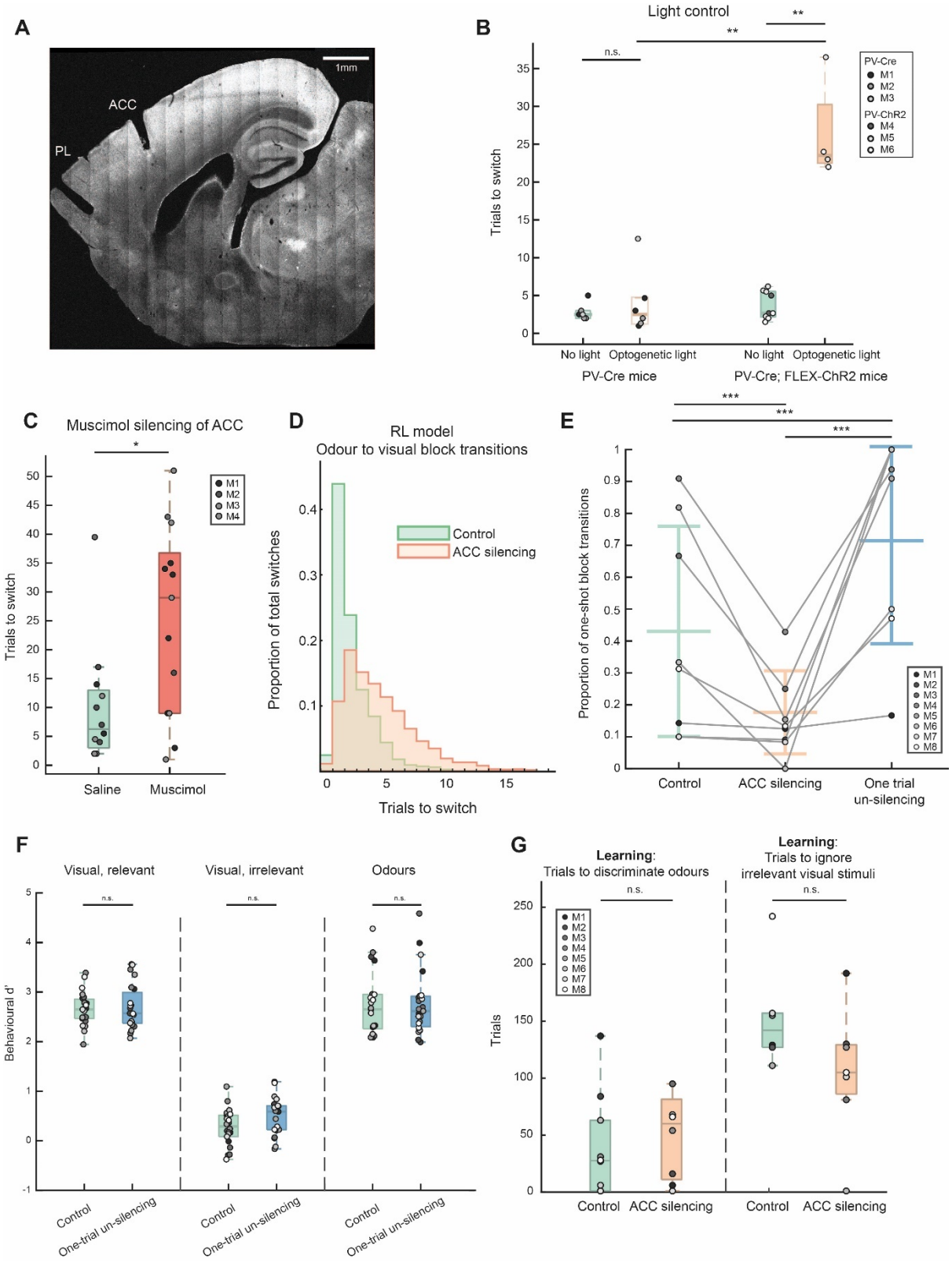
1252 Behavioural d-primes in each silencing condition were calculated by taking performance in stable
1253 periods of blocks, outside of transition periods [74,75]. Switching speeds were defined as the number of
1254 trials that elapse before a mouse correctly responded to three rewarded visual gratings in a row, either
1255 licking to the grating after a switch to a visual block or by ignoring the grating after a switch to an
1256 odour block. A 'fluke' switch was defined as a switch in which the mouse correctly licked to the very
1257 first rewarded grating in a new visual block, before any evidence of the switch had been received.
1258 These were interpreted as exploratory licks, visible in the histograms in Figure 1F and Supplementary
1259 Figure 3E at trial 0, and all such transitions were excluded from analysis of switching speeds.

## SUPPLEMENTARY FIGURES



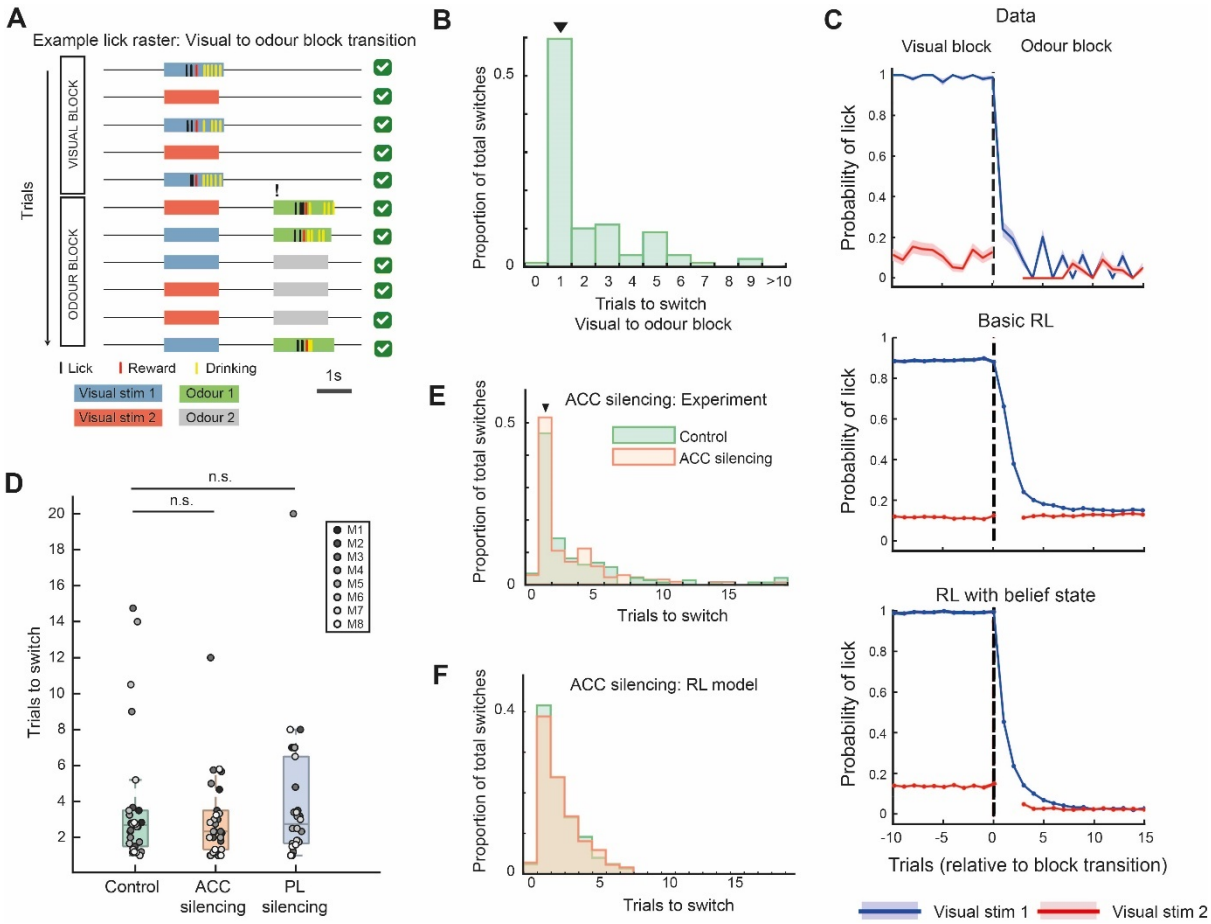



**Supplementary Figure 1 | Mice perform rapid task-switching with one-shot block transitions. A.** Schematic showing default stimulus durations used in the task, with representative licking and reward valve signals. Top, visual stimulus 1 (rewarded) and 2 (unrewarded) visual block trials. Bottom, example odour 1 (rewarded) or odour 2 (unrewarded) odour block trials. **B**. 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.

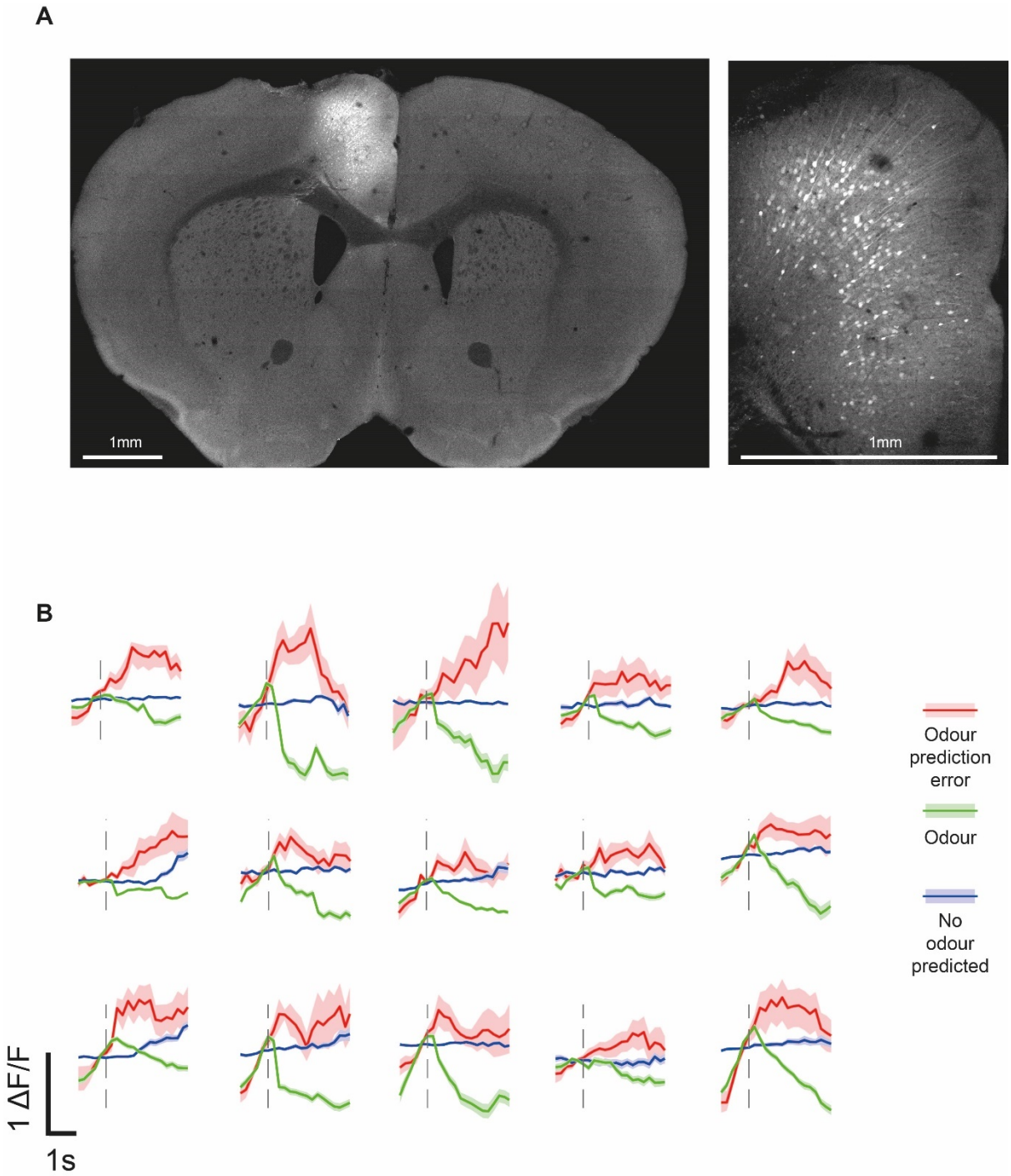



**Supplementary Figure 2 | ACC inhibition impairs task switching. A.** Sagittal slice from experimental mouse showing relative placements of ACC and PL targeting optic fibre cannulae. **B.** Average number of trials required in a session to switch from an odour block to a visual block, N = 3 PV-Cre (6 sessions per condition) and 3 PV-Cre;FLEX-Channelrhodopsin2 (9 no light and 4 light sessions) mice, median ± IQR here and below. PV-Cre mice: no light 3 ± 1.21, optogenetic light 3 ± 4.12. PV-Cre;FLEX-ChR2 mice: no light 3 ± 4.56, optogenetic light 23 ± 8.65. Wilcoxon rank sum test comparing PV-Cre mice in light

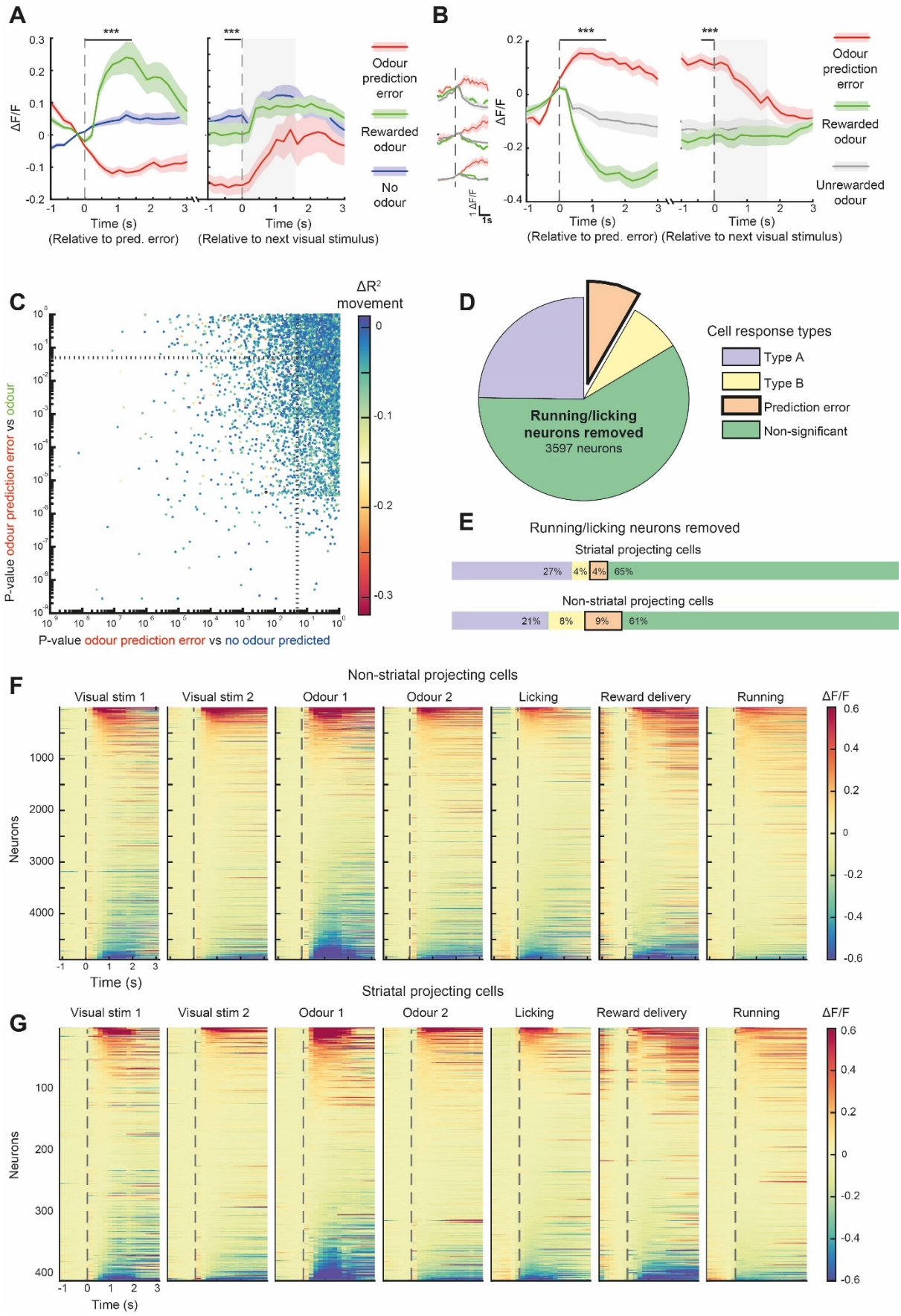
and no-light conditions P = 0.857, comparing PV-Cre and PV-Cre;ChR2 mice in light conditions ** P = 0.0095, comparing PV-Cre;ChR2 mice in light and no light conditions ** P = 0.0028. **C.** Average number of trials per session required to switch from an odour block to a visual block, following bilateral infusions of saline or muscimol into the ACC. Saline infusions, 13 sessions, 6.25 ± 10, muscimol infusions, 14 sessions, 29 ± 27.75, N = 4 mice. Wilcoxon rank-sum comparison P = 0.03. **D.** 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 0.21. **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.** Steady-state task discrimination performance after each successful block transition. Behavioural d-primes to visual stimuli in a visual block, median ± IQR: control and one trial un-silencing 2.65 ± 0.39 and 2.57 ± 0.62 respectively, visual stimuli in an odour block 0.29 ± 0.43 and 0.59 ± 0.48, and odour stimuli 2.65 ± 0.69 and 2.62 ± 0.61. Wilcoxon signed-rank test comparing d-primes in control and one trial un-silencing sessions to visual stimuli in visual block P = 0.9757, to visual stimuli in odour block P = 0.0883, to odour stimuli P = 0.9861. **G.** 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.




**Supplementary Figure 3. Visual to odour 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. **B.** Histogram showing the number of trials required to switch from visual to odour block, with one-shot transitions indicated by arrowheads. 115 transitions respectively, 17 sessions, 14 mice. **C.** 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. Data from 115 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. **D.** 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. **E.** Histogram of number of trials taken to switch between blocks during control and continuous ACC silencing sessions. Transitions from visual to odour block, N = 148 and 146 control and ACC silencing transitions respectively. One-shot transitions indicated by arrowhead. **F.** RL model with belief state fit to the data, with and without ACC silenced, as in Supplementary Figure 2D. Number of trials taken by the agent to switch from visual to odour blocks did not change significantly. The model fit the data best when the amplitude of the mismatch signal was reduced by a factor of 0.5.

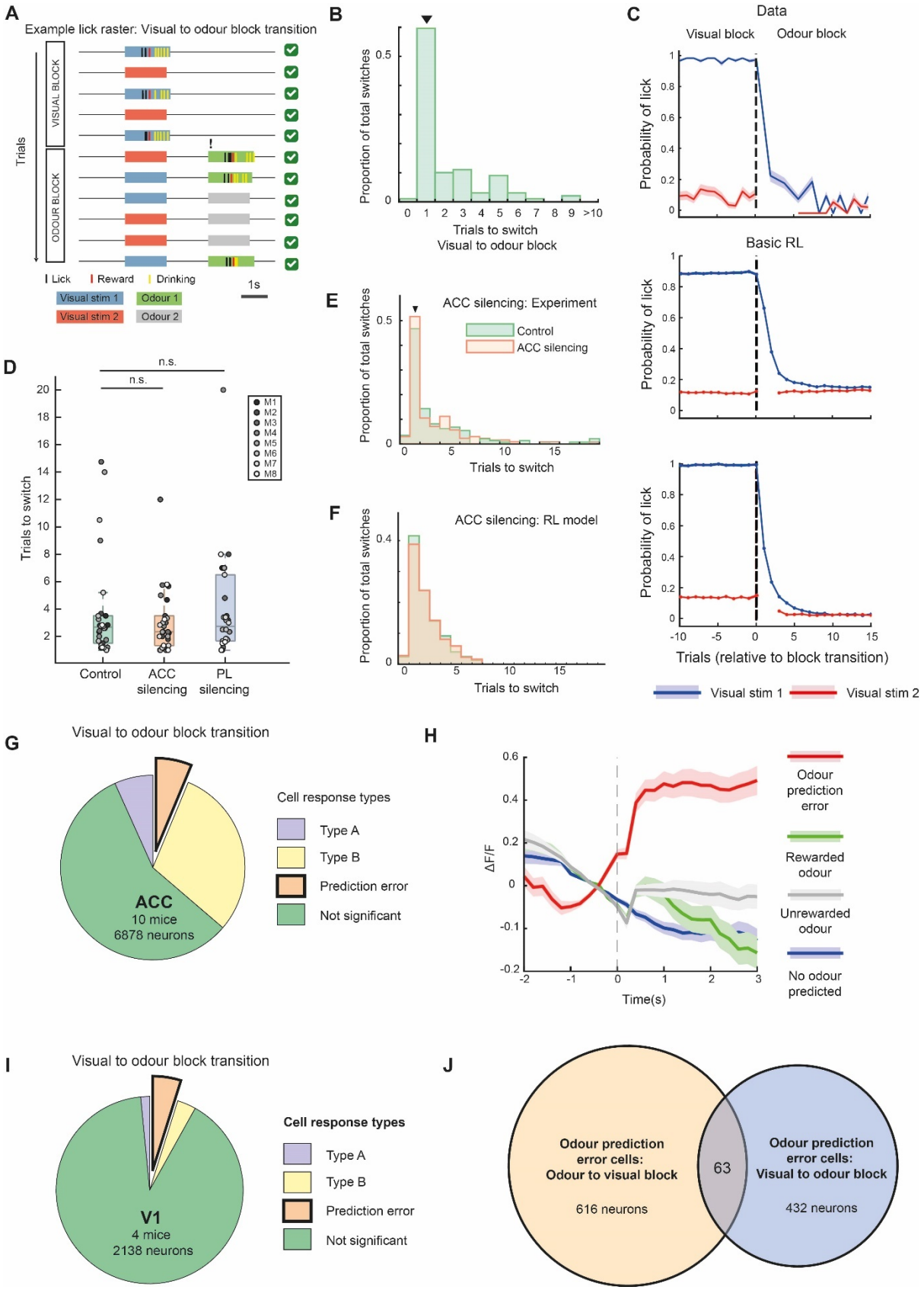



**Supplementary Figure 4 | ACC neurons respond to prediction errors. A.** Example coronal slice (~+1.2mm anterior relative to bregma) showing GCaMP7f expression relative to the whole brain (left) and within the ACC (right). **B.** Example prediction-error neurons identified from ACC recordings in the 10 imaged mice, showing averaged responses in the odour prediction error, odour, and no odour predicted trials, -1 to 3s relative to real or predicted odour onset.



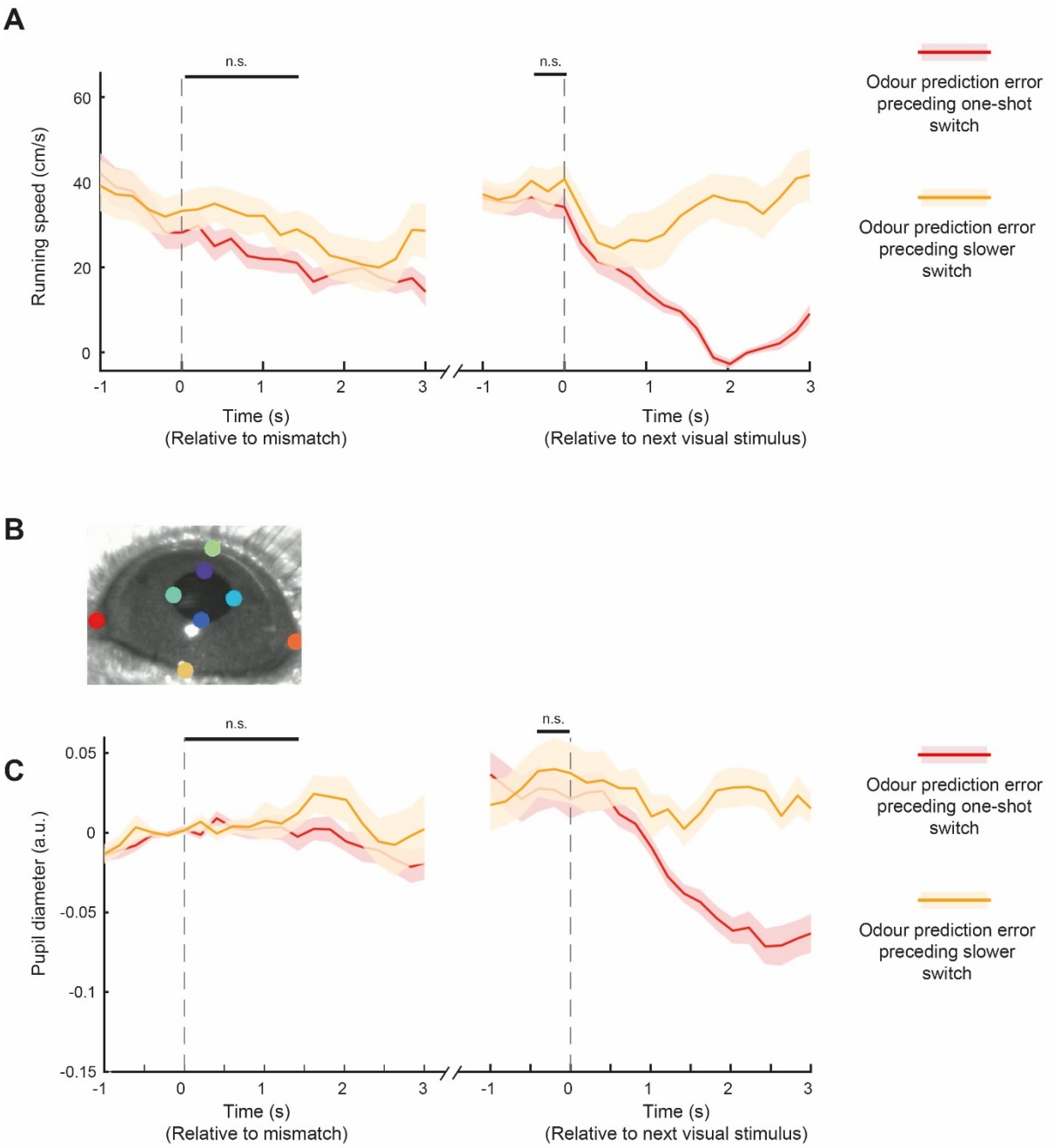

**Supplementary Figure 5 | Response properties of ACC neurons. A.** Mean response of all negatively responding prediction-error neurons (N = 186 cells) aligned to the odour prediction error event, and

to the next visual stimulus onset (grey shading). Response shown to the odour prediction error event (red) and actual odour delivery (green). Wilcoxon signed-rank test, prediction-error response vs baseline averaged 0 to 1.5 s from prediction error 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, prediction error response vs odour response, *** $P = 2.97 \times 10^{-19}$, $P = 1.26 \times 10^{-6}$ respectively. **B.** Mean response of all positively responding prediction-error neurons (N = 168 cells) aligned to odour prediction error event, and to the next visual stimulus onsets (grey shading), shown with responses to the odour prediction error event (red) and both the actual rewarded odour (green) and unrewarded odour (grey) delivery. Wilcoxon signed-rank test between the odour prediction error response and unrewarded odour response averaged 0 to 1.5s from prediction error event and -0.5 to 0.s from the next stimulus onset, *** $P = 1.69 \times 10^{-28}$, $P = 5.45 \times 10^{-15}$. **C.** The influence of running and licking on each neuron was 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 neuron (N = 6878 cells). Prediction-error neurons had significantly different activity between the prediction error vs odour, and prediction error vs no odour conditions (lines for $P < .05$ shown on each axis). Neurons with stronger influence of running and licking (more negative $\Delta R^2$) were not enriched in prediction-error neurons. **D.** Proportions of cells as in Figure 4h, after removing cells which were influenced by running (N = 3597 cells). **E.** Proportions of cells as in Figure 4l, after removing cells which were influenced by running (N = 202 and 2437 striatal projecting and non-striatal projecting cells respectively, same colour legend as **D**). **F.** 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. **G.** As in **F**, but for striatal projecting neurons (N = 421 cells).




**Supplementary Figure 6 | Prediction-error neurons in visual to odour block transitions. A.** Proportions of cells, as in Figure 4h, for the visual to odour block transition. N = 463 Type A, 6.7%, 2063 Type B, 29.9%, and 432 prediction-error neurons, 6.2%, total 6878 cells, 10 mice. **B.** Mean response of all positive prediction-error neurons (N = 432 cells) aligned to the odour prediction error

event (arrival of an unexpected odour). Response shown to the odour prediction error 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 **A**, for primary visual cortex (V1). N = 33 Type A, 1.5%, 69 Type B, 3.2%, and 104 prediction-error neurons, 4.8%, total 2138 cells, 4 mice. **D**. Venn diagram showing the overlap between prediction-error neurons in odour to visual block transitions and prediction-error neurons in visual to odour block transitions. The proportion of overlapping cells is significantly higher than chance (bootstrap test, $P < 0.01$).




**Supplementary Figure 7 | Running and pupil size during one-shot and slower prediction errors. A.** Averaged running speed during odour prediction-error trials preceding either one-shot or slower switches, or odour trials. Wilcoxon signed-rank test comparing running speeds in odour prediction-error trials preceding one-shot and slower switches, aligned to expected odour onset 0 to 1.5s P = 0.101, aligned to the next visual stimulus (grey shading) -0.5 to 0s P = 0.413. **B.** Example frame from eye camera recording, with coloured markers used to track eye width/height and pupil width/height using DeepLabCut. **C.** Averaged pupil area during odour prediction-error trials preceding either one-shot or slower switches, or odour trials. Wilcoxon signed-rank test comparing pupil area in odour prediction-error trials preceding one-shot and slower switches, aligned to expected odour onset 0 to 1.5s P = 0.734, aligned to the next visual stimulus (grey shading) -0.5 to 0s P = 0.4258.