Thoughtful faces: inferring internal states across species using facial features

Alejandro Tlaie[1,2,✉], Muad Y. Abd El Hay[1], Berkutay Mert[1], Robert Taylor[1], Pierre-Antoine Ferracci[1], Katharine Shapcott[1], Mina Glukhova[1], Jonathan W Pillow[3], Martha N Havenith[1,*], and Marieke Schölvinck[1,*]

[1]Ernst Strüngmann Institute for Neuroscience in cooperation with the Max Planck Society, Frankfurt am Main, 60528, Germany
[2]Laboratory for Clinical Neuroscience, Centre for Biomedical Technology, Universidad Politécnica de Madrid, Spain
[3]Princeton Neuroscience Institute, Princeton University, Princeton, NJ 08544, USA
[*]These authors contributed equally.

Animal behaviour is shaped to a large degree by internal cognitive states, but it is unknown whether these states are similar across species. To address this question, we developed a virtual reality setup in which mice and macaques engage in the same naturalistic visual foraging task. We exploited the richness of a wide range of facial features extracted from video recordings during the task, to train a Markov-Switching Linear Regression (MSLR). By doing so, we identified, on a single-trial basis, a set of internal states that reliably predicted when the animals were going to react. At any one moment, one state was dominant, and mice transitioned through these dominant states faster than monkeys. The model could predict not only reaction time, but also task outcome, supporting the behavioural relevance of the inferred states. The thus characterised states were similar between mice and monkeys. Thus, we were able to flexibly and agnostically track the dynamics of several internal states, identifying general principles of naturalistic cognitive processing across species.

Internal states | Naturalistic behavior | Cross-species
Correspondence: ***atboria@gmail.com***

Introduction

In the wild, all mammals show similar behaviour: they all hunt or forage for food, sleep, mate, avoid predators, and explore their environment, to name just a few. None of these behaviours can simply be explained as a passive reaction to environmental input - rather, they are crucially shaped by dynamic fluctuations in internal states such as satiety, alertness, curiosity or attention (1, 2). So, if fundamental behaviours are comparable across species, how similar are the internal states that drive them? Is 'attention' in a monkey the same as 'attention' in a mouse?

The common approach towards investigating internal states has been a reductionist one: highly restrictive tasks featuring simplified stimuli and requiring narrow behavioural repertoires (e.g. button presses), with little room for fluctuations over time (3–5). What's more, experimental paradigms diverge widely depending on the species under study. For example, attention studies in primates typically require the subject to fixate on a central fixation point while paying attention to a peripheral stimulus that might briefly or subtly change its appearance (6, 7). Attention studies in rodents, on the other hand, typically use the 5-choice serial reaction time task (5CSRTT), in which the subject is required to scan a row of five apertures for the presentation of a brief light stimulus, and then navigate towards the light source (8, 9). Even though the behaviour associated with high attention, i.e. short reaction times and accurate responses, is the same in both cases, clearly these tasks are too different to draw any meaningful cross-species comparisons.

Breaking away from this restrictive regime towards studying internal states as they occur naturally is challenging. To tackle this challenge successfully, an ideal behavioural paradigm needs to (1) rely on innate, naturalistic behaviours to accurately reflect spontaneously occurring rather than training-induced internal states (10), (2) identify internal states in a data-driven way that is not restrained by (potentially anthropomorphising) concepts of cognitive processing imposed by the researcher, and (3) track the evolution of internal states over time to capture their intrinsically dynamic nature. For this, binary metrics of behaviour such as a button presses or nose pokes will not suffice; rather, precise, multi-parametric behavioural tracking is needed to generate time-resolved analyses that extract the evolution of underlying cognitive states from the measured behavioural parameters moment by moment (11–13).

Recent technological advances have opened up new avenues to achieve these goals in a principled way. Virtual reality (VR) environments, for instance, allow researchers to create immersive yet highly controlled experimental settings that can be tailored to different species' intrinsic sensory capacities and behavioural repertoires (14, 15). For instance, color spectra can easily be adapted to the visibility range of most species; and input devices like treadmills allow animals to interface with the VR in largely intuitive ways that require minimal training (16). As such, VR provides a powerful tool for studying animal behaviour in a highly controlled yet dynamic and ethologically sound context. Importantly, this maximizes adaptability across species, opening up the unique opportunity to record directly comparable behaviours in different species.

At the same time, advances in deep-learning algorithms enable us to dynamically track ongoing changes in body movement and posture from video footage (17–19). These algorithms have opened up new avenues to harvest an unprecedented amount of information even from simple behavioural paradigms, far exceeding classical behavioural readouts such as button presses or saccades (20, 21). Most importantly, they allow for the ongoing and time-resolved tracking of be-

havioural dynamics - a fundamental prerequisite if we aim to identify the spontaneous emergence of internal cognitive and emotional states (22, 23).

In this study, we leverage these technological breakthroughs to infer and directly compare the internal states of two species commonly studied in neuroscience - macaques and mice. Specifically, we combine a highly immersive and naturalistic VR foraging task with a state-of-the-art deep learning tool that allows for precise, automated tracking of behavioural features. The features extracted in this way then serve as inputs to a Markov-Switching Linear Regression (MSLR) model (24), which finally predicts the ongoing probabilities of different internal states on a single-trial basis.

Importantly, such single-trial inference of internal states is only meaningful if the behavioural markers it relies on are not indirectly tracking the concrete motor outputs required for task performance. If the behaviour was directly linked to task-related motor output (e.g. preparatory paw movements), then internal states inferred from this behaviour would be expected to trivially predict task performance. To ensure that the behavioural parameters we chose would truly reflect internal processing, we focused on the animals' facial expressions.

While facial expressions have long been thought to only play a role in highly visual and social species like monkeys and humans (25–28), recent work has highlighted that also less social, less visual species like mice exhibit meaningful facial expressions (22, 29). As such, behaviourally relevant facial expressions seem to be much more evolutionarily preserved than previously expected (22, 29, 30). More specifically, they seem to reflect fundamental emotions like pleasure, pain, disgust and fear in a way that is not only consistent within one species, but also readily translatable across species (31, 32). This argues for an evolutionary convergent role of facial expressions in reflecting (and potentially communicating) emotions.

Unlike these previous studies on the relation between facial expressions and emotions, here we for the first time analyse facial expressions in mice and monkeys that occur spontaneously, in the absence of a pre-defined emotional context. Such spontaneously occurring behavioural states have so far mainly been tracked using single facial features to track isolated cognitive states, for instance by tracking attention via pupil size, both in rodents (33, 34) and primates (35–38). Similarly, eye movements in monkeys and humans (39–41) and whisker movements in mice (42) have been used to track attention and decision-making. By focusing on entire facial expressions beyond individual (often species-specific) features, we aim to for the first time map out the spectrum of spontaneously occurring internal states in a way that is 1) agnostic, i.e. not focused on a specific cognitive process, and 2) directly comparable across species.

Our approach of using facial expressions to infer internal states from natural behaviour constitutes a drastic move away from the classical approach of imposing internal states through restrictive behavioural paradigms (e.g. cued attentional shifts). By tying the results of this approach back to known relationships between internal states and overt behaviour, such as shorter reaction times during focused attention, these data-driven, agnostically inferred internal states can be tentatively related to known cognitive processes such as attention and motivation. Importantly, this puts us in the unique position to directly compare inferred internal states across two species.

Results

A. Experimental set-up. To track and compare spontaneously occurring internal states of mice and macaques during the performance of the same naturalistic visual discrimination task, the animals were placed inside a custom-made spherical dome (Fig. 1A, top). On the inside of the dome, we projected a virtual reality (VR) environment using a custom-made toolbox called DomeVR (16). The monkeys navigated through the VR environment manually using a trackball; the mice ran on a spherical treadmill, the movements of which were translated into VR movements (for details, see Methods - *Experimental Setup*).

Two monkeys and seven mice were used in this study, comprising 18 and 29 sessions respectively. The animals engaged in a simple, foraging-based two-choice perceptual decision task, in which they had to approach a target stimulus while avoiding a distractor stimulus, both of which were represented by natural leaf shapes integrated in a meadow landscape (Fig. 1A, bottom; see Methods - *Behavioral paradigm* and *Behavioral Training*). Their performance on this task was quantified first in terms of trial outcomes: hit (target stimulus reached), wrong (distractor stimulus reached), and miss (neither stimulus reached); as well as in reaction time (RT). For this, we identified turning points in the animals' running trajectories through the VR to define the moment when an animal decisively oriented itself towards one of the two potential targets (Fig. 1C; for details, see Methods - *Reaction Time*). As Fig. S1 shows, success rate and reaction times were largely comparable across species, although mice showed less consistent performance than monkeys, in terms of running trajectories, reaction times, and correct target choices. We hypothesize that this is due to the lack of fine motor control of the mice on the trackball.

As the animals were performing the task, we recorded their faces. For macaques, this was done by analysing video footage from one camera positioned frontally on the monkey's face, as well as eye tracking output (see Methods - *Behavioural tracking*). For mice, we analysed video footage from one camera positioned on the side of the face (Fig. 1B). From these videos, we extracted facial features such as eyebrow, nose and ear movement using DeepLabCut (Fig. 1C; see Methods - *Facial key point extraction*). For monkeys, we selected 18 features; for mice, 9 features (see Methods - *Face features* for the full list of facial features).

For each trial, facial features were averaged over a time window of $250ms$ before the stimuli appeared in the VR environment. This time window was chosen to maximize the interpretability of the inferred hidden states: as there is no task-relevant information available yet, presumably all of the

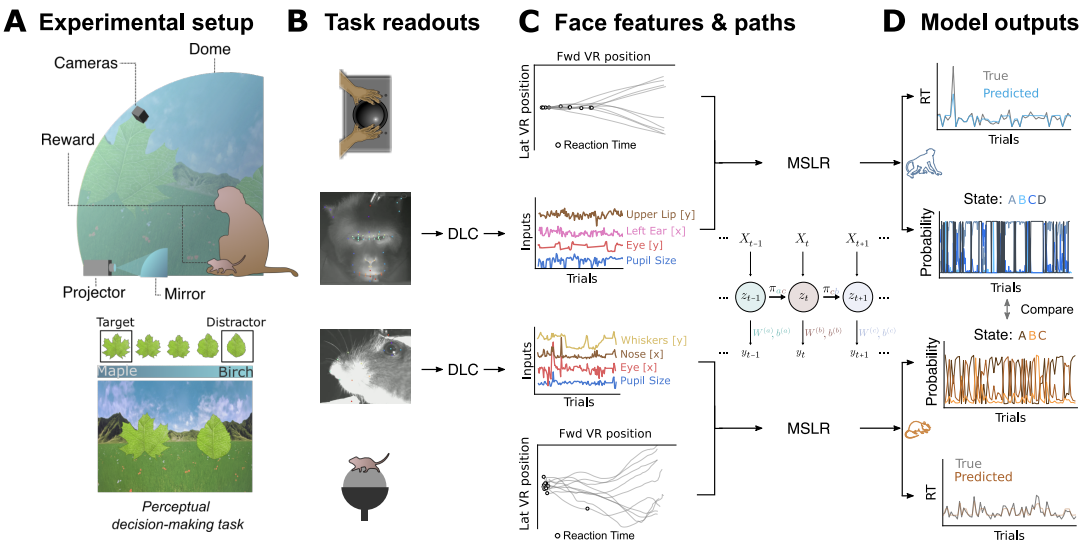

Fig. 1. Experimental setup and computational pipeline. A) Macaques and mice were seated inside a large dome on the inside of which a VR was projected via a curved mirror (top). They were rewarded for moving towards a spike-shaped leaf compared to a round-shaped leaf (bottom). **B)** As the animals were engaged in the task, behavioural data were collected: movements of the trackball (top and bottom) and videos of their faces (middle). **C)** Trackball movements were translated into paths through the virtual environment (top and bottom), from which reaction times were determined (see Methods). Individual facial features were automatically detected from the videos and tracked over time (middle). **D)** Facial features entered two separate MSLR models (one for each species), which yielded, for every trial, a predicted reaction time and internal state probabilities.

facial expressions that the animal makes are due to internally generated processes, rather than being reflective of stimulus or task parameters.

B. Model performance. The facial features extracted in this way were used as inputs to a Markov-Switching Linear Regression (MSLR) model (Fig. 1C; see Methods - *MSLR*). The MSLR manages to reflect the non-stationarity and regime shifts often present in behavioural data (43–46), by flexibly accommodating complex temporal dynamics while keeping a relative simplicity, compared to deep learning-based methods (47). Moreover, the MSLR is less data-hungry than other common data-driven models (48–50). The MSLR uses the 'pre-stimulus' facial features to predict the animals' reaction time (RT) by assuming 'hidden' states. Each hidden state implies a different relation between individual facial features and the subsequent RT in the same trial. For instance, in one hidden state, the RT might be best predicted by eyebrow movements, while in another, nose sniffing might be most predictive. The overall number of hidden states is determined by the experimenter (see below). For each trial, the model then outputs the predicted RT as well as the probability of each hidden state (Fig. 1D). The two models (one for mice, one for monkeys) were trained and tested on data from all individuals; S14 shows the outcomes of the models split by session and by individual.

To test if this approach was appropriate for our behavioural recordings, we first checked if assuming the presence of multiple hidden states was in fact warranted by the data, or if they could also be described by one constant, uniform relationship between facial expressions and RTs over time. To this end, we determined model performance when only one internal state was permitted (Fig. 2A). For both species, the model's predictive performance was remarkably low under these circumstances - in fact, predictions were less accurate than random guessing.

Next, we quantified model performance for different numbers of hidden states - which is the main free parameter of the MSLR. Model performance was tested by using cross-validation (see Methods - *Model tuning*). For both species, the cross-validated R^2 improved dramatically when allowing for more than one hidden state until reaching a plateau. Since the accuracy of RT predictions began to saturate with increased model complexity, we took the finite difference of the CV performance curve for each species and fixed the number of internal states at its maximum, in order to reach the optimal trade-off between predictive accuracy and model simplicity (Fig. 2A). This approach arrived at a similar optimal number of hidden states for both species: For monkeys, the optimal number of states was 4, for mice it was 3. Tests on held-out data showed a similar performance (Fig. 2A, insets), indicating that the high predictive performance was not due to overfitting.

In both species, our models yielded remarkably accurate trial-by-trial predictions of RT, indicating that pre-trial facial expressions can indeed predict subsequent task performance (Fig. 2B, top row). It also suggests that the relation between facial features and task performance is dynamic rather than static over time, reflecting multiple underlying states.

The coexistence of several hidden states opens up the question if task performance is dominated by a single state at any given moment, or if several states co-exist continuously. The probabilities of each state over time suggest the latter (Fig. 2B, bottom row). These observations were confirmed by the highly bimodal distribution of these probabilities for both species (Fig. 2C). This indicates that trials typically feature one dominant state (probability close to one) and several largely absent states (probability close to zero). In monkeys,

this separation between dominant and non-dominant states was particularly pronounced, while in mice, state probabilities were somewhat more mixed. Quantifying the state dominance through the Kullback-Leibler divergence (KL) corroborated these findings (Fig. 2D). So, the hidden states identified by our model seem to reflect mostly mutually exclusive behavioural modes that animals switch in and out of. Given how consistently trials were dominated by one state, we chose to binarize hidden state outcomes by assigning each trial to its most probable hidden state.

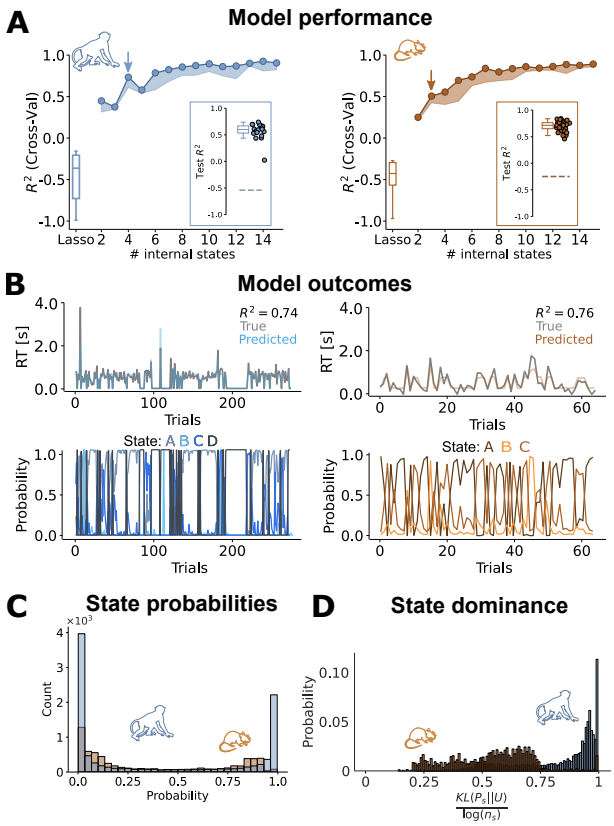

Fig. 2. Model performance and state probabilities. A) Cross-validation performance for various numbers of states, for macaques (left) and mice (right). Circles indicate the maximum CV R^2 and the shaded region extends until the 5th percentile. For both species, increasing the number of states improves model performance to a plateau at an $R^2 \approx 0.8$. *Lasso* is a regularized Linear Regression (i.e., a MSLR with 1 internal state). The arrows indicate the number of states we selected, based on the maximum difference of the CV performance curve (see Fig S10). Insets show model performance for held out data at the selected number of states; dashed horizontal lines indicate the 99th percentile of the surrogate performances (see Methods). Note that the shuffled R^2 is negative, because only uncorrelated predictors are expected to be centered at 0, and due to finite sampling effects, there is always a non-zero correlation between the shuffling and the ground-truth. Furthermore, as we are dealing with skewed distributions (see Fig. S1), the null tendency is not captured by the mean, as assumed by the default R^2. **B)** Predicted RTs (top) and state probabilities (bottom) for an example stretch of data (left, macaques; right; mice). **C)** Probabilities of all states over all trials, regardless of state identity (blue, macaques; orange, mice). The bimodal distribution suggests that states are either absent or dominant on any given trial. **D)** Kullback-Leibler divergence (KL) for monkey (blue) and mouse (orange) internal states. KL quantifies the difference between the state probabilities and the uniform distribution, normalizing by the number of states. A KL value close to 1 indicates maximally dissimilar distributions (i.e., only one present state at a time), while a value close to 0 indicates indistinguishable distributions (i.e., equally likely states).

C. State dynamics. To explore if the hidden states showed attributes that could be reflective of internal cognitive states, we first characterized their temporal dynamics. To this end, we examined the frequency of state transitions in both species. The state transition matrices, which show how likely a trial of a given hidden state is followed by a trial of any (other or same) state (Methods - *Markov-Switching Linear Regression*), revealed high values along the diagonal for macaques. This indicates that their hidden states are highly stable (Fig. 3A, left). In contrast, in mice state transitions were generally more frequent, with just one state (state A) slightly more self-sustaining than the others (Fig. 3A, right). As a complementary analysis, we computed the dwell time for each state. This quantity is defined as the number of consecutive trials that a given state is occupied for, before transitioning to a different state. Hidden states lasted up to three times longer in macaques than in mice, suggesting that internal processing may be more steady in macaques (Fig. 3B). This is consistent with previous findings that behavioural dynamics may fluctuate faster in mice (34, 51) than monkeys (52), and may be a sign of the well trained monkeys having a more stereotyped behavioural strategy for solving the task than the lesser trained mice.

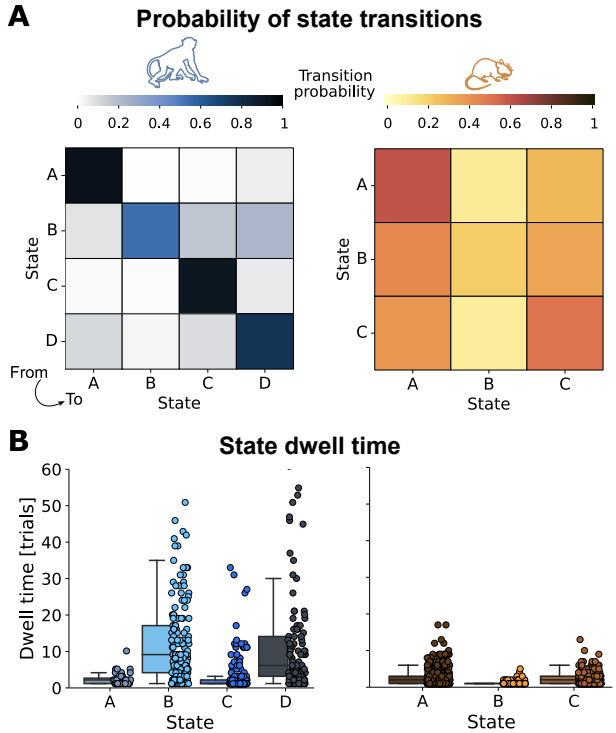

Fig. 3. State dynamics. A) State transition matrices for macaques (left) and mice (right), that show the probability, at any one trial, of transitioning from a certain state (rows) to any other state (columns). Transitions between different states (off-diagonal terms) are more frequent for mice than for macaques. **B)** Macaques (left) spend more time than mice (right) in the same state, as measured by the dwell time (number of consecutive trials of each state being the most likely one).

D. Hidden states as performance states. To link the identified hidden states more concretely to internal cognitive processing, we set out to investigate how each hidden state related to behavioural outcomes, starting with the RTs that the model was trained to predict. There are two potential scenarios for this: On the one hand, it is possible that each hidden state covers the full range of RTs, but predicts them from a

different constellation of facial features. Alternatively, each hidden state might 'specialize' on predicting specific ranges of RTs. For example, one hidden state might cover facial features that distinguish between fast and extremely fast RTs, while another state mainly predicts variations between slower RTs. This second scenario would make it more likely that the identified hidden states reflect genuinely distinct performance states.

To distinguish between these scenarios, we plotted the overall state-specific RT distributions, pooling trials across all sessions and animals, for each hidden state (Fig. 4A; Fig. S15 shows the same plot for individual sessions and animals). The resulting distributions mainly support the second scenario: While one hidden state (state B in both monkeys and mice) covered a rather broad range of RTs, all other states showed a distinct profile of response speeds. This implies that the hidden states relate to distinct performance regimes (in this case in terms of response speed), making them viable candidates for defining specific internal states of cognitive task processing.

To further probe the possible link of our internal states to known cognitive processes, we related all hidden states to the three possible trial outcomes of the task (hit, wrong, and miss; see Methods - *Task performance and internal states*). Crucially, given that we trained the model to predict RTs, it never received any explicit information about trial outcome. Furthermore, RTs were only marginally related to trial outcomes (Fig. S1), so that trials with a specific RT would not be significantly more likely to result e.g. in a hit or a miss trial. Finally, as we only used information about facial features in the pre-stimulus phase of the trial to train the model, it cannot reflect stimulus features.

Even though information about trial outcomes was not part of the MSLR model, the resulting hidden states were consistently predictive of a specific trial outcome (Fig. 4B). For instance, in monkeys, trials that were classified as belonging to state C were most likely to result in a hit, while trials from state A often resulted in incorrect responses, even though the RT distributions of both states overlapped strongly. For both animals, different states were significantly correlated with different behavioral outcomes.

Combining these effects of internal states on RTs and outcome revealed specific combinations of speed and accuracy. We plotted mean RT per hidden state against the probability of hit trials in the same state. The constellation of states in this space was noticeably similar across species (Fig. 4C). Both mouse and monkey data seem to generate a hidden state (State C) that is associated with relatively fast RTs and largely successful trial outcomes - a performance regime that could be interpreted as globally attentive. Conversely, state A in both species reflects fast but often incorrect responses, potentially reflecting more impulsive decision making (53). Finally, state B for both species features particularly slow RTs and mostly misses for mice and equally likely trial outcomes for monkeys; this could signify a state of global inattention. The only state that appears in monkeys but not mice (state D) features no reactions at all (i.e. no change in path direction) and only misses; a sign of complete task disengagement.

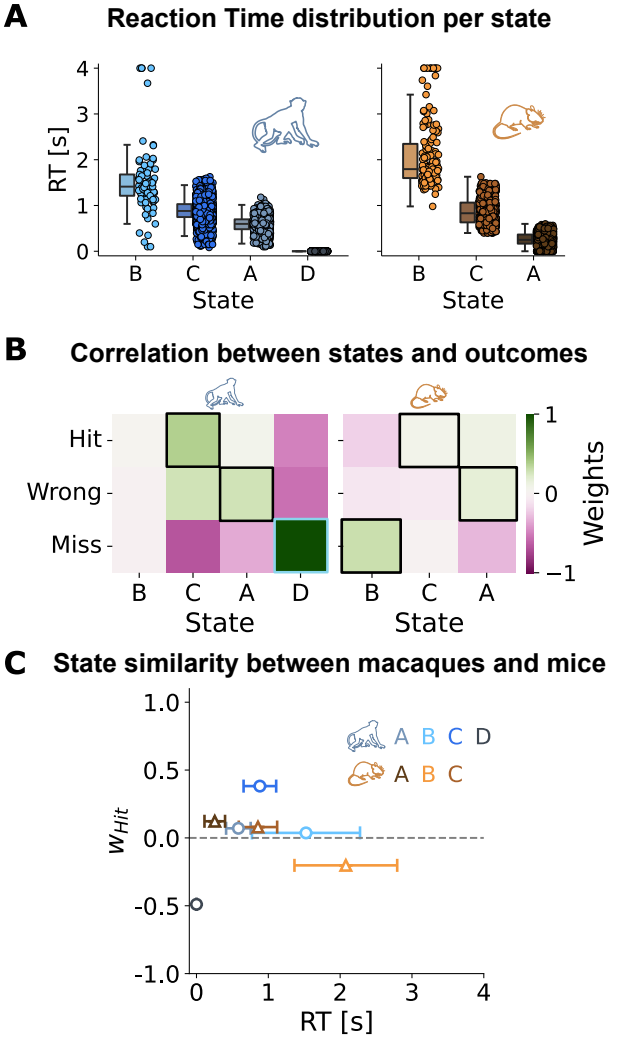

Fig. 4. Internal states and task performance. A) Splitting the RTs over internal states shows large diversity for both macaques (left) and mice (right), from fast reaction-states to extremely slow ones. **B)** Correlations of state probabilities with the three task outcomes (hit, wrong, miss), for macaques (left) and mice (right). Black boxes indicate the states most strongly associated with a certain task outcome. **C)** Conjunction of RT and probability of a hit outcome, for all states (blue circles, macaque; orange triangles, mouse).

E. Relationship to facial features. A final clue towards the interpretation of our internal states might be given by the facial features from which they are inferred. To this end, we plotted the regression weights of the facial features for the hidden states associated with hit, wrong, and miss trials (Fig. 5A; for the facial features comprising the fourth state in the monkey, see Fig. S8).

These plots reveal highly varying contributions of facial features to the internal states. For example, in mice, pupil size is only predictive of fast RTs during hit trials, but not in wrong or miss trials. In monkeys, the predictive role of pupil size is reversed in miss trials compared to others, and ear movements play a strong predictive role in hit and wrong, but not miss trials. The variability of all model-input facial features across the various states confirms that in both species, there is not just one facial feature that distinguishes between

the states, but that all features contribute (Fig 5B). Together, these results suggest that 1) holistic analysis of facial features is much more informative than selecting single features and 2) the relationship between facial features and cognitive processing is not linear, but changes depending on the state the animal is in.

Interestingly, the facial expressions predicting RTs in hit and wrong trials are quite similar in monkeys but not mice. This resonates with the similar link of these two states to the trial outcomes (Fig 4B), potentially suggesting that both happen with high attention, and wrong trials are mainly attributed to visual difficulty (see Fig. S13 for a summarized visualization).

One reason why hidden states can predict trial outcomes so accurately despite not being trained on them in any way might be that pre-trial facial features are mostly a trivial consequence of the animal's trial history. For example, facial features might mainly reflect an animal still drinking reward from the previous trial, which might in turn raise motivation to perform correctly in the upcoming trial. In this case, facial features would merely be a particularly convoluted way of quantifying the previous trial outcome, and using it to predict upcoming performance, as has been achieved previously (REFs Carandini, other foraging papers??). To account for this possibility, we trained an Auto-Regressive Hidden Markov Model (ARHMM) based on RTs (see Methods - *ARHMM* for details). As can be seen in Fig. S5, the facial features model outperforms the ARHMM for all states, for both species.

As an extra control, we correlated each facial feature with the history of prominent task parameters, specifically two related to the directly previous trial (its outcome, which might affect motivation; and the location of its target, which might predict side biases), and two related to the overall session history (the cumulative amount of reward and the time that passed since the start of the session, as proxies for satiety and fatigue, respectively). Correlations between task variables and facial features were sparse in both species (Fig. S11). In fact, attributes of the previous trial did not relate significantly to facial features at all, and more sustained session attributes modulated facial features merely somewhat. This suggests that facial features may be modulated by ubiquitous internal processes like fatigue and satiety, which are in turn impacted by task parameters, but they are not a trivial reflection of task history. Rather, the fact that facial expressions are modulated by the overall task context makes them a more plausible reflection of realistic fluctuations in cognitive processing.

Discussion

Internal cognitive states are known to substantially shape overall brain activity (54, 55) as well as behavioural decision making (1, 2), yet they are notoriously difficult to identify. As a result, it is even less clear to what extent they converge across species. To infer hidden cognitive states in mice and monkeys, we harnessed an MSLR model (24) trained on their facial features while they were engaged in an immersive VR foraging task. Specifically, we trained the MSLR to predict an animal's reaction time (RT) in a given trial based on its facial expressions prior to stimulus presentation. For both species, RTs could be predicted with high accuracy from preceding facial features only, suggesting that facial expressions reflect parameters that are directly relevant to task performance. These parameters were only minimally shaped by task history, suggesting that they were not a trivial reflection e.g. of the previous trial outcome.

Even more surprisingly, this approach revealed multiple distinct hidden states, which were characterized by equally distinct relationships between facial features and task performance. For instance in monkeys, in one hidden state, reaction times were best predicted by pre-stimulus ear and right eye brow movements, while in another state, pupil size and the left eye brow dominated the reaction time prediction in terms of predictive power. Thus, depending on an animal's internal state, the relation between facial features and subsequent task performance can shift dramatically.

These findings stand in marked contrast to previous research, which has mainly highlighted linear relationships between single facial features (e.g. pupil size or eye movements) and isolated cognitive states (e.g. attention) (33–41). Our findings imply that such analyses miss out on a large portion of the information available through complex facial expressions. Our findings were not dependent on the use of one specific MSLR model, as we repeated our analyses with a different model and training pipeline (a GLM-HMM, training an individual model per animal and experimental session), with very similar results (56).

Most importantly, the internal states revealed in this manner mapped robustly onto behavioural trial outcomes (i.e. hit, wrong and miss trials) - even though this information had been in no way part of the inputs the MSLR received. This suggests that the hidden states highlighted by the MSLR were not simply 'computational devices' increasing the model's predictive power. Instead, they appear to reflect genuine, dynamically fluctuating cognitive states, which result in distinctive behavioural outcome profiles.

Interestingly, despite the fact that the optimal number of states was determined separately for each species and in a purely data-driven way, our approach converged onto a low and noticeably similar number of internal states for both species: three states for mice, four for macaques. How comparable are these internal states of mice and monkeys?

We found that in terms of the dynamics by which animals traversed different internal states, results diverged strongly across species. Specifically, mice transitioned more frequently between states than monkeys. A control analysis that matched the number of subjects, trials and facial parameters across species before fitting the MSLR models showed that this difference is not a trivial result of divergences in data structure (see Suppl Fig 11). Given that mice have previously been shown to alternate between strategies during perceptual decision-making (43), this finding may point at a genuine difference in the cognitive dynamics of mice and monkeys. Alternatively, the prolonged training time of the monkeys over the mice may have given them the chance to converge on

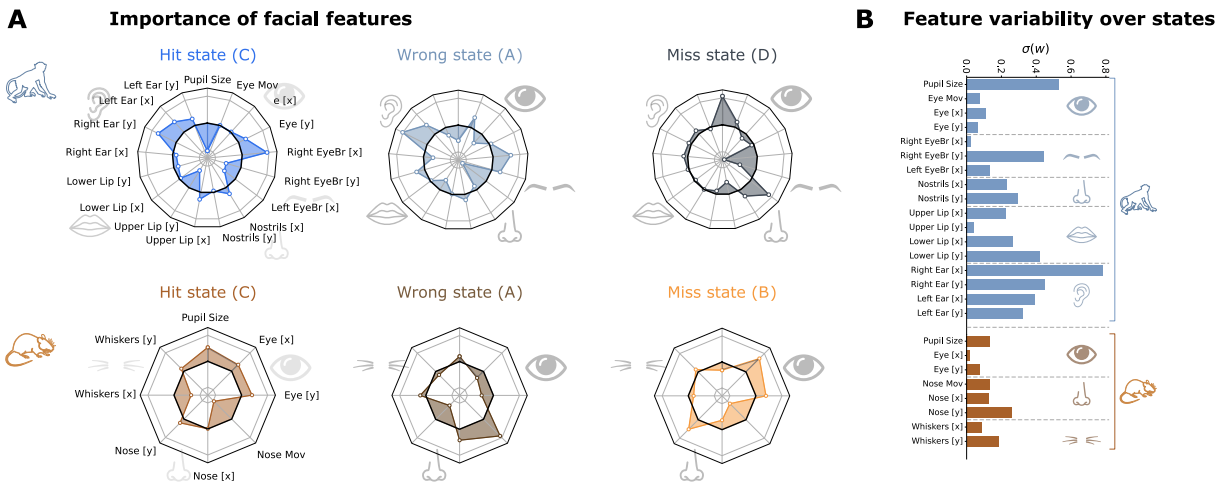

Fig. 5. Informativeness of facial features A) Predictor weights of the facial features for the macaque (top) and mouse (bottom) model in the hit, wrong and miss states (see black boxes in Fig4B). Central circle indicates a predictor weight of zero, inside this circle are negative predictor weights, outside are positive weights. Each state has its own characteristic facial expression pattern. **B)** Variability of all facial features over states. Although some features contribute more than others, clearly all features contribute to the model distinguishing between the various internal states.

more stable behavioural strategies over the course of training (57, 58). Recordings from mice that experienced a more prolonged training scheme and/or from more naive monkeys will give fascinating insights into the role of expertise in fostering more stable transitions between cognitive states.

Despite these differences in internal state dynamics, the constellation of behavioural profiles covered by different states was remarkably similar across species. Each hidden state predicted only a narrow range of reaction times; and when relating the inferred hidden states to task performance beyond the RTs that the model was trained to predict, we found that the states mapped onto the various behavioural outcomes (hit, wrong, miss) with varying probabilities. Moreover, each hidden state covered unique combinations of RT ranges and trial outcomes (hit, miss and wrong trials), despite the fact that trial outcomes had not been part of the MSLR in any way.

The constellation of states in this task performance space was highly comparable across species. (59, 60). Specifically, both monkeys and mice display a state where trial outcome is typically slow and incorrect (which could be interpreted as 'inattentive'), as well as several states where performance is largely fast and correct, with a preference for thoroughness in one state, and a preference for speed in the other. In monkeys, one additional species-specific state predicted no reactions and a mixture of all possible outcomes, indicating complete disengagement from the task.

These states potentially map on various levels of task-related attention, and further support the notion that classical concepts of attention do indeed reflect much of the internal structure of goal-directed behaviour, also in naturalistic settings. Together, these results suggest that, rather than serving as mere 'computational prediction devices', the hidden states inferred agnostically by the MSLR map onto genuine and readily interpretable cognitive states, which are structured in surprisingly similar ways across species.

The fact that different states are associated with distinct constellations of facial features points to a role of facial expressions beyond emotional expression. Facial expressions have so far been mostly studied in a social or emotional context, and mostly in social species such as monkeys (28, 61) and humans (25, 62). In mice, until recently facial expressions were thought to mainly reflect pain (29, 63, 64), until careful analyses using machine-learning algorithms identified their facial expressions as innate and accurate reflections of several emotional states as well (22, 31, 65). Our results suggest that similarly to humans, facial expressions in monkeys and mice also convey cognitive and motivational variables such as focus or cognitive strain, even in the absence of a particular emotional or social context.

The fact that such performance-related states are equally apparent in both species is particularly surprising, since one can assume that the prominent differences between the two species (such as the acuity and dominance of their visual system) would imply that they are likely to solve tasks using different strategies. Yet both species are subject to task-independent internal states, such as slow fluctuations in attention (52–54). We believe that the internal states we tap into with the current approach are more reminiscent of such fluctuations, and as such, can be important indicators of underlying brain-wide activity fluctuations. Global internal states such as arousal, motivation, and attention typically manifest themselves via brain-wide dynamics, and it will be an exciting endeavour to investigate how well those neuronally defined internal states correspond to the ones we here identified behaviourally.

The MSLR model that we used yielded single estimates of the internal states per trial. This constitutes a great basis for time-resolved tracking of internal states, which can be further extended in future, using MSLR models with higher temporal resolution. Such MLSR models will be able to align identified internal states with specific events within each trial, such as the appearance and disappearance of stimuli, thereby allowing for more precise characterisation of their dynamics and functional roles.

Perhaps even more importantly, such a time resolved MSLR

would also allow us to link cognitive processes to neural activity on a moment by moment basis, without the need for repeatedly presenting identical trials and then doing extensive post-hoc averaging. As the MSLR model yields a time-resolved estimate of cognitive states, these time courses can be directly compared to continuous neural activity. As such, this approach opens up a much more naturalistic view of the neuro-behavioural dynamics involved in spontaneous cognitive states than traditional approaches can offer (44, 66).

These findings suggest that in an ecologically valid framework that applies across species (in this case, a foraging-based task set in a naturalistic, immersive visual environment), many features of cognitive processing are more similar than classical paradigms might have suggested. At the same time, presumably genuine cross-species differences, e.g. in the transition frequency between cognitive states, also become more apparent.

In summary, we have shown here that in both monkeys and mice, facial features can be used to infer internal cognitive states, and to track their spontaneous dynamics over time. With this approach, we find that the basic attributes of such internal states map onto known cognitive states such as attention in both species in a highly translatable way, but that the dynamic by which mice and monkeys traverse these different states is clearly different. This highlights the crucial importance of using naturalistic behavioural paradigms, especially in cross-species research, in order to discern truly species-specific results from differences induced by restrictive testing methods.

ACKNOWLEDGEMENTS

We thank Carlos Wert Carvajal for useful input on earlier versions of the manuscript. We thank Victor Geadah, Iris Stone and Matthew Creamer for their valuable comments on model training. A.T. acknowledges support from the Margarita Salas Fellowship (Spanish Ministry of Economy) and from the the Add-On Fellowship for Interdisciplinary Life Sciences (Joachim Herz Stiftung).

Bibliography

1. Steven W Flavell, Nadine Gogolla, Matthew Lovett-Barron, and Moriel Zelikowsky. The emergence and influence of internal states. *Neuron*, 2022.
2. David Deutsch, Diego Pacheco, Lucas Encarnacion-Rivera, Talmo Pereira, Ramie Fathy, Jan Clemens, Cyrille Girardin, Adam Calhoun, Elise Ireland, Austin Burke, et al. The neural basis for a persistent internal state in drosophila females. *Elife*, 9:e59502, 2020.
3. John W Krakauer, Asif A Ghazanfar, Alex Gomez-Marin, Malcolm A MacIver, and David Poeppel. Neuroscience needs behavior: correcting a reductionist bias. *Neuron*, 93(3): 480–490, 2017.
4. Matteo Grasso, Larissa Albantakis, Jonathan P Lang, and Giulio Tononi. Causal reductionism and causal structures. *Nature neuroscience*, 24(10):1348–1355, 2021.
5. Gaël Varoquaux and Russell A Poldrack. Predictive models avoid excessive reductionism in cognitive neuroimaging. *Current opinion in neurobiology*, 55:1–6, 2019.
6. Conrado A Bosman, Jan-Mathijs Schoffelen, Nicolas Brunet, Robert Oostenveld, Andre M Bastos, Thilo Womelsdorf, Birthe Rubehn, Thomas Stieglitz, Peter De Weerd, and Pascal Fries. Attentional stimulus selection through selective synchronization between monkey visual areas. *Neuron*, 75(5):875–888, 2012.
7. Georgia G Gregoriou, Stephen J Gotts, Huihui Zhou, and Robert Desimone. High-frequency, long-range coupling between prefrontal and visual cortex during attention. *science*, 324(5931):1207–1210, 2009.
8. Trevor Robbins. The 5-choice serial reaction time task: behavioural pharmacology and functional neurochemistry. *Psychopharmacology*, 163:362–380, 2002.
9. Samuel K Asinof and Tracie A Paine. The 5-choice serial reaction time task: a task of attention and impulse control for rodents. *JoVE (Journal of Visualized Experiments)*, (90): e51574, 2014.
10. Ann Kennedy. The what, how, and why of naturalistic behavior. *Current opinion in neurobiology*, 74:102549, 2022.
11. Alex Gomez-Marin, Joseph J Paton, Adam R Kampff, Rui M Costa, and Zachary F Mainen. Big behavioral data: psychology, ethology and the foundations of neuroscience. *Nature neuroscience*, 17(11):1455–1462, 2014.
12. Alex Gomez-Marin and Asif A Ghazanfar. The life of behavior. *Neuron*, 104(1):25–36, 2019.
13. Yael Niv. The primacy of behavioral research for understanding the brain. *Behavioral Neuroscience*, 135(5):601, 2021.
14. Christopher D Harvey, Philip Coen, and David W Tank. Choice-specific sequences in parietal cortex during a virtual-navigation decision task. *Nature*, 484(7392):62–68, 2012.
15. Roberto A Gulli, Lyndon R Duong, Benjamin W Corrigan, Guillaume Doucet, Sylvain Williams, Stefano Fusi, and Julio C Martinez-Trujillo. Context-dependent representations of objects and space in the primate hippocampus during virtual navigation. *Nature neuroscience*, 23(1):103–112, 2020.
16. Katharine A Shapcott, Marvin Weigand, Iuliia Glukhova, Martha N Havenith, and Marieke L Schölvinck. Domevr: A setup for experimental control of an immersive dome virtual environment created with unreal engine 4. *[Manuscript submitted for publication]*, 2022. doi: 10.1101/2022.04.04.486889.
17. Alexander Mathis, Pranav Mamidanna, Kevin M Cury, Taiga Abe, Venkatesh N Murthy, Mackenzie Weygandt Mathis, and Matthias Bethge. Deeplabcut: markerless pose estimation of user-defined body parts with deep learning. *Nature neuroscience*, 21(9):1281–1289, 2018.
18. Alexander B Wiltschko, Tatsuya Tsukahara, Ayman Zeine, Rockwell Anyoha, Winthrop F Gillis, Jeffrey E Markowitz, Ralph E Peterson, Jesse Katon, Matthew J Johnson, and Sandeep Robert Datta. Revealing the structure of pharmacobehavioral space through motion sequencing. *Nature neuroscience*, 23(11):1433–1443, 2020.
19. Kevin Luxem, Petra Mocellin, Falko Fuhrmann, Johannes Kürsch, Stephanie R Miller, Jorge J Palop, Stefan Remy, and Pavol Bauer. Identifying behavioral structure from deep variational embeddings of animal motion. *Communications Biology*, 5(1):1267, 2022.
20. Michael Pereira, Pierre Megevand, Mi Xue Tan, Wenwen Chang, Shuo Wang, Ali Rezai, Margitta Seeck, Marco Corniola, Shahan Momjian, Fosco Bernasconi, et al. Evidence accumulation relates to perceptual consciousness and monitoring. *Nature communications*, 12(1):3261, 2021.
21. Michael H McCullough and Geoffrey J Goodhill. Unsupervised quantification of naturalistic animal behaviors for gaining insight into the brain. *Current Opinion in Neurobiology*, 70: 89–100, 2021.
22. Nate Dolensek, Daniel A Gehrlach, Alexandra S Klein, and Nadine Gogolla. Facial expressions of emotion states and their neuronal correlates in mice. *Science*, 368(6486):89–94, 2020.
23. Sandeep Robert Datta, David J Anderson, Kristin Branson, Pietro Perona, and Andrew Leifer. Computational neuroethology: a call to action. *Neuron*, 104(1):11–24, 2019.
24. Richard E Quandt. A new approach to estimating switching regressions. *Journal of the American statistical association*, 67(338):306–310, 1972.
25. Paul Ekman and Harriet Oster. Facial expressions of emotion. *Annual review of psychology*, 30(1):527–554, 1979.
26. Ralph Adolphs, Hanna Damasio, Daniel Tranel, and Antonio R Damasio. Cortical systems for the recognition of emotion in facial expressions. *Journal of neuroscience*, 16(23):7678–7687, 1996.
27. Robert Aubrey Hinde and Thelma E Rowell. Communication by postures and facial expressions in the rhesus monkey (macaca mulatta). In *Proceedings of the Zoological Society of London*, volume 138, pages 1–21. Wiley-Blackwell, 1962.
28. Lisa A Parr and Matthew Heintz. Facial expression recognition in rhesus monkeys, macaca mulatta. *Animal behaviour*, 77(6):1507–1513, 2009.
29. Dale J Langford, Andrea L Bailey, Mona Lisa Chanda, Sarah E Clarke, Tanya E Drummond, Stephanie Echols, Sarah Glick, Joelle Ingrao, Tammy Klassen-Ross, Michael L LaCroix-Fralish, et al. Coding of facial expressions of pain in the laboratory mouse. *Nature methods*, 7(6):447–449, 2010.
30. Christian L Ebbesen and Robert C Froemke. Body language signals for rodent social communication. *Current opinion in neurobiology*, 68:91–106, 2021.
31. Olivia Le Moëne and Max Larsson. A new tool for quantifying mouse facial expressions. *Eneuro*, 10(2), 2023.
32. Anna D Zych and Nadine Gogolla. Expressions of emotions across species. *Current opinion in neurobiology*, 68:57–66, 2021.
33. Dan Alin Ganea, Alexander Bexter, Mathias Günther, Pierre-Marie Gardères, Björn M Kampa, and Florent Haiss. Pupillary dilations of mice performing a vibrotactile discrimination task reflect task engagement and response confidence. *Frontiers in Behavioral Neuroscience*, 14:159, 2020.
34. Jacob Reimer, Emmanouil Froudarakis, Cathryn R Cadwell, Dimitri Yatsenko, George H Denfield, and Andreas S Tolias. Pupil fluctuations track fast switching of cortical states during quiet wakefulness. *neuron*, 84(2):355–362, 2014.
35. Paola Binda, Maria Pereverzeva, and Scott O Murray. Pupil size reflects the focus of feature-based attention. *Journal of neurophysiology*, 112(12):3046–3052, 2014.
36. R Becket Ebitz, John M Pearson, and Michael L Platt. Pupil size and social vigilance in rhesus macaques. *Frontiers in neuroscience*, 8:100, 2014.
37. Chin-An Wang and Douglas P Munoz. A circuit for pupil orienting responses: implications for cognitive modulation of pupil size. *Current opinion in neurobiology*, 33:134–140, 2015.
38. Siddhartha Joshi and Joshua I Gold. Pupil size as a window on neural substrates of cognition. *Trends in cognitive sciences*, 24(6):466–480, 2020.
39. Tirin Moore and Mazyar Fallah. Control of eye movements and spatial attention. *Proceedings of the National Academy of Sciences*, 98(3):1273–1276, 2001.
40. Burkhart Fischer and B Breitmeyer. Mechanisms of visual attention revealed by saccadic eye movements. *Neuropsychologia*, 25(1):73–83, 1987.
41. Miriam Spering. Eye movements as a window into decision-making. *Annual review of vision science*, 8:427–448, 2022.
42. Sina E Dominiak, Mostafa A Nashaat, Keisuke Sehara, Hatem Oraby, Matthew E Larkum, and Robert NS Sachdev. Whisking asymmetry signals motor preparation and the behavioral state of mice. *Journal of Neuroscience*, 39(49):9818–9830, 2019.
43. Zoe C Ashwood, Nicholas A Roy, Iris R Stone, International Brain Laboratory, Anne E Urai, Anne K Churchland, Alexandre Pouget, and Jonathan W Pillow. Mice alternate between discrete strategies during perceptual decision-making. *Nature Neuroscience*, 25(2):201–212, 2022.
44. Luca Mazzucato. Neural mechanisms underlying the temporal organization of naturalistic

animal behavior. *Elife*, 11:e76577, 2022.
45. Adam J Calhoun, Jonathan W Pillow, and Mala Murthy. Unsupervised identification of the internal states that shape natural behavior. *Nature neuroscience*, 22(12):2040–2049, 2019.
46. Daniel Hulsey, Kevin Zumwalt, Luca Mazzucato, David A McCormick, and Santiago Jaramillo. Decision-making dynamics are predicted by arousal and uninstructed movements. *bioRxiv*, pages 2023–03, 2023.
47. Pedro J Gonçalves, Jan-Matthis Lueckmann, Michael Deistler, Marcel Nonnenmacher, Kaan Öcal, Giacomo Bassetto, Chaitanya Chintaluri, William F Podlaski, Sara A Haddad, Tim P Vogels, et al. Training deep neural density estimators to identify mechanistic models of neural dynamics. *Elife*, 9:e56261, 2020.
48. Byron M Yu, John P Cunningham, Gopal Santhanam, Stephen Ryu, Krishna V Shenoy, and Maneesh Sahani. Gaussian-process factor analysis for low-dimensional single-trial analysis of neural population activity. *Advances in neural information processing systems*, 21, 2008.
49. Mohammad Reza Keshtkaran, Andrew R Sedler, Raeed H Chowdhury, Raghav Tandon, Diya Basrai, Sarah L Nguyen, Hansem Sohn, Mehrdad Jazayeri, Lee E Miller, and Chethan Pandarinath. A large-scale neural network training framework for generalized estimation of single-trial population dynamics. *Nature Methods*, 19(12):1572–1577, 2022.
50. Sean M Perkins, John P Cunningham, Qi Wang, and Mark M Churchland. Simple decoding of behavior from a complicated neural manifold. *bioRxiv*, pages 2023–04, 2023.
51. James FA Poulet and Sylvain Crochet. The cortical states of wakefulness. *Frontiers in systems neuroscience*, 12:64, 2019.
52. Marlene R Cohen and John HR Maunsell. Using neuronal populations to study the mechanisms underlying spatial and feature attention. *Neuron*, 70(6):1192–1204, 2011.
53. Marieke N Havenith, Petra M Zijderveld, Stefan van Heukelum, Rogier B Poorthuis, Huibert D Mansvelder, and August B Smit. The virtual-environment-foraging task enables rapid training and single-trial metrics of attention in head-fixed mice. *Scientific Reports*, 8(1): 17371, 2018. doi: 10.1038/s41598-018-34966-8.
54. Kenneth D Harris and Alexander Thiele. Cortical state and attention. *Nature reviews neuroscience*, 12(9):509–523, 2011.
55. Marieke L Schölvinck, Aman B Saleem, Andrea Benucci, Kenneth D Harris, and Matteo Carandini. Cortical state determines global variability and correlations in visual cortex. *Journal of Neuroscience*, 35(1):170–178, 2015.
56. Alejandro Tlaie, Katharine Shapcott, Muad Abd el Hay, Berkutay Mert, Pierre-Antoine Ferracci, Robert Taylor, Iuliia Glukhova, Martha N Havenith, and Marieke Schölvinck. Thoughtful faces: Using facial features to infer naturalistic cognitive processing across species. Cosyne Abstracts, 2023. doi: 10.1145/1217935.1217951.
57. Patricia D Stokes. Learned variability. *Animal Learning & Behavior*, 23:164–176, 1995.
58. Patricia D Stokes, Francis Mechner, and Peter D Balsam. Effects of different acquisition procedures on response variability. *Animal Learning & Behavior*, 27:28–41, 1999.
59. Ben Mitchinson and Tony J Prescott. Whisker movements reveal spatial attention: a unified computational model of active sensing control in the rat. *PLoS computational biology*, 9(9): e1003236, 2013.
60. Ehud Fonio, Goren Gordon, Noy Barak, Yonatan Winetraub, Tess Baker Oram, Sebastian Haidarliu, Tali Kimchi, and Ehud Ahissar. Coordination of sniffing and whisking depends on the mode of interaction with the environment. *Israel Journal of Ecology & Evolution*, 61(2): 95–105, 2015.
61. Xin-He Liu, Lu Gan, Zhi-Ting Zhang, Pan-Ke Yu, and Ji Dai. Probing the processing of facial expressions in monkeys via time perception and eye tracking. *Zoological Research*, 44(5): 882, 2023.
62. Alan J Fridlund. *Human facial expression: An evolutionary view.* Academic press, 2014.
63. Alexander H Tuttle, Mark J Molinaro, Jasmine F Jethwa, Susana G Sotocinal, Juan C Prieto, Martin A Styner, Jeffrey S Mogil, and Mark J Zylka. A deep neural network to assess spontaneous pain from mouse facial expressions. *Molecular pain*, 14:1744806918763658, 2018.
64. Amy L Miller and Matthew C Leach. The mouse grimace scale: a clinically useful tool? *PloS one*, 10(9):e0136000, 2015.
65. Erwin B Defensor, Michael J Corley, Robert J Blanchard, and D Caroline Blanchard. Facial expressions of mice in aggressive and fearful contexts. *Physiology & behavior*, 107(5): 680–685, 2012.
66. Simon Musall, Matthew T Kaufman, Ashley L Juavinett, Steven Gluf, and Anne K Churchland. Single-trial neural dynamics are dominated by richly varied movements. *Nature neuroscience*, 22(10):1677–1686, 2019.
67. Keiji Matsuda, Takeshi Nagami, Yasuko Sugase, Aya Takemura, and Kenji Kawano. A widely applicable real-time mono/binocular eye tracking system using a high frame-rate digital camera. In *Human-Computer Interaction. User Interface Design, Development and Multimodality: 19th International Conference, HCI International 2017, Vancouver, BC, Canada, July 9-14, 2017, Proceedings, Part I 19*, pages 593–608. Springer, 2017.
68. Shaokai Ye, Anastasiia Filippova, Jessy Lauer, Maxime Vidal, Steffen Schneider, Tian Qiu, Alexander Mathis, and Mackenzie Weygandt Mathis. Superanimal models pretrained for plug-and-play analysis of animal behavior, 2023.
69. Chang. P, Harper-Donnelly. G, Kara A, Li X, et al, and Linderman S Murphy K. Dynamax, a library for State Space Models. *GitHub*, 2022.
70. Arthur P Dempster, Nan M Laird, and Donald B Rubin. Maximum likelihood from incomplete data via the em algorithm. *Journal of the royal statistical society: series B (methodological)*, 39(1):1–22, 1977.
71. Scott Linderman, Matthew Johnson, Andrew Miller, Ryan Adams, David Blei, and Liam Paninski. Bayesian learning and inference in recurrent switching linear dynamical systems. In *Artificial Intelligence and Statistics*, pages 914–922. PMLR, 2017.
72. Takuya Akiba, Shotaro Sano, Toshihiko Yanase, Takeru Ohta, and Masanori Koyama. Optuna: A next-generation hyperparameter optimization framework. In *Proceedings of the 25th ACM SIGKDD international conference on knowledge discovery & data mining*, pages 2623–2631, 2019.
73. Tom AB Snijders. On cross-validation for predictor evaluation in time series. In *On Model Uncertainty and its Statistical Implications: Proceedings of a Workshop, Held in Groningen, The Netherlands, September 25–26, 1986*, pages 56–69. Springer, 1988.
74. Nikolaus Hansen. The cma evolution strategy: A tutorial. *arXiv preprint arXiv:1604.00772*, 2016.

Methods

Animals. This study includes data from two male macaques (*Macaca mulatta*) and six male Black6 mice (*Mus musculus*). All procedures were approved by the regional authorities (*Regierungspräsidium Darmstadt*) under the authorization number F149/2000 and were performed in accordance with the German law for the protection of animals and the "European Union's Directive 2010/63/EU".

Surgical Procedures. All animals were fitted with custom-milled headposts for the purpose of head fixation during this experiment. The headpost design and implant procedures for the macaques have been extensively discussed by Psarou et al (2023). Briefly, a four-legged titanium baseplate was screwed into the skull under general anesthesia. After several weeks of osseo-integration, a titanium top part was screwed onto the baseplate in a simple procedure. The headposts for the mice have been described here (JoVE article). Briefly, the animal was placed under isoflurane anaesthesia, shaved and given local analgesia on the top of the head. An incision was made and the skin on top of the cranium was removed, before the cranium was cleaned and the custom milled titanium head plate was attached using dental cement.

Experimental Setup. Experiments were carried out in a darkened room (mice) or electrically shielded booth (monkeys). The animals were in the centre of a 120-cm diameter spherical dome extending to 250 deg visual angle. The headfixed mice were positioned on a styrofoam spherical treadmill; the headfixed monkeys were seated in a monkey chair and could spin a 12-cm diameter trackball with their hands. Movements of the spherical treadmill and trackball allowed the animals to traverse a virtual reality (VR) environment projected on the inside of the dome by means of a spherical mirror. Projecting the VR environment on a dome surrounding the animals enabled both their central and peripheral view to be covered, thereby providing an immersive and realistic VR environment. The VR environment was created using DomeVR, our custom-made toolbox combining photorealistic graphics rendered with Unreal Engine 4, with high timing precision required for neuroscience experiments ((16)).

Experimental Paradigm. Mice and monkeys were required to distinguish two natural shapes at equal distance in front of them, amidst a grassy field with a blue sky above and mountains in the background (Fig 1A). The two shapes emerged out of a central shape which was either right at the starting position (for monkeys) or a short distance in front (for mice). A virtual collision with the correct shape yielded a reward ('hit'), whereas the incorrect shape yielded no reward ('wrong'), and no collision with either shape also yielded no reward ('miss') (2AFC paradigm). Rewards were drops of diluted juice for the monkeys and drops of vanilla soy milk for the mice. For the monkeys, the shapes varied smoothly between a non-rewarded, textured square and a rewarded triangle (monkey K) or between a rewarded, jagged and a non-rewarded, hour-glass shaped leaf (monkey C). On each trial, a blend between the two shapes was shown alongside the exact middle blend ('reference shape'). For the mice, the shapes and their reward contingencies were the same as for monkey C.
Monkey data were recorded in xx sessions for monkey C, yy sessions for monkey K. Each session lasted about one hour, during which the monkeys completed xx +- yy trials at xx and yy percent correct (monkeys C and K, respectively). The monkeys were both fully trained on handling the trackball to move through the VR environment, as well as the VR task. Following the headpost surgery, the mice were handled for 5 days to reduce experimental anxiety due to head fixation and interaction with the experimenter, before behavioural training began. Behavioural training in the experimental setup at initial stages lasted between 3-5 sessions, before final data collection began, which lasted up to 30 sessions.

Behavioural tracking. We recorded videos of the monkeys' and mice' faces during the tasks at 60 Hz using Basler acA640-121gm infrared cameras with a modified version of PylonRecorder2 software (https://gitlab.mpcdf.mpg.de/mpibr/scic/pylonrecorder/PylonRecorder2). Additionally, in the monkeys, eye movements were recorded at 500 Hz using a Grasshopper3 infrared camera and the free eye tracking software *iRecHS2* (67) and synchronized with DomeVR (16).

Facial key point extraction. To extract facial key points from the videos, we used markerless pose estimation on them, as implemented in DeepLabCut (17, 68). For mice, features were extracted from videos of the left side of the face using our own model to identify key points such as the coordinates of the eye, whisker pad and nose. For mouse pupillometry, we used the eye coordinates from the face model to crop the video to include the entire left eye and ran it through a refined model based on the "mouse pupil vclose" Animal Zoo model (provided by Jim McBurney-Lin at the University of California Riverside, USA) included with DeepLabCut. The output of the pupil model was 8 points covering the circumference of the mouse pupil, that were then used to calculate pupil and eye summary statistics.
For the macaque facial key points, we used the pre-trained "primate face" model from the DeepLabCut Animal Zoo (provided by Claire Witham at the Centre for Macaques, MRC Harwell, UK) and extended it with additional points on the lips to capture more precise mouth movement than in the original model. All models were further trained and refined to achieve a detection error of less than 2 pixels per tracked key point in all conditions. The macaque raw pupil size recorded by the eye-tracker was Z-scored over time within the training data set.

To synchronise the video timing with events in the virtual reality environment, we used 32 ms long infrared flashes emitted from an LED mounted near the camera lens. These flashes were then extracted from the face videos to be used as timestamps for synchronisation with DomeVR. Five consecutive flashes indicated the start of a behavioural session; a single flash indicated the start of a trial.

Reaction Time. In our VR setting, where animals move towards one of two stimuli rather than pressing a button or lever, or making an eye movement, we define the reaction time (RT) as the time point of the initial substantial movement directed towards either stimulus. While determining this time point, it is crucial to distinguish between stimulus-related movements and minor positional adjustments. We specifically focus on the first deviation in lateral movement, while excluding forward movement due to its susceptibility to random movements and its task irrelevance.

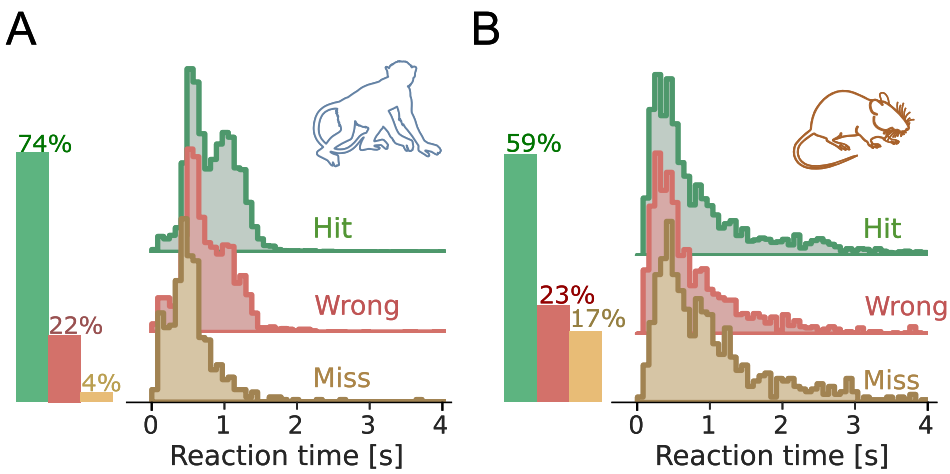

Fig. S1. Reaction Times. Distribution of Reaction Times for macaques (A) and mice (B), split by behavioral outcome; data are pooled over sessions ($n = 18$ and $n = 28$ for macaques and mice, respectively). The three distributions largely overlap.

To calculate the RT, we use a sliding window linear regression approach, incorporating a time decay mechanism. This approach enables us to detect non-linearity by examining the coefficient of determination (R^2) for each window. A low R^2 value indicates that the data deviate from linearity, and such a deviation can be interpreted as a deviation in lateral movement.
First we compute a linear regression on the time series of lateral VR movement for adjacent sliding windows i and j of a given size (n_w). Then, R_i^2 (i.e., R^2 for window i) is calculated as:

$$R_i^2 = 1 - \frac{\sum_{j=1}^{n_w} \left(l_j - \hat{l}_j^i\right)^2}{\sum_{j=1}^{n_w} \left(l_j - \bar{l}\right)^2} \tag{1}$$

where l_j is the jth element of the lateral movement observed in the second window, $\hat{l}_j^i$ is the corresponding predicted lateral movement value (based on window i) and $\bar{l}$ is the mean lateral movement within the second window. As a result, we get an array of R^2 values: $\mathbf{R}^2 = [R_1^2, \ldots, R_n^2]$.
Subsequently, we reverse the sign of the $-\mathbf{R}^2$ array and detect its local maxima. For this, we resort to the definition of extreme points (we have a univariate function in this case):

$$L = \underset{w}{\operatorname{argmax}} \left[\frac{d^2 r(w)}{dw^2} \right] \tag{2}$$

where we have simplified the notation, using $-\mathbf{R}^2 \equiv r(w)$. Once we have found the local maxima (L), we further require that they have a minimum prominence (λ). Prominence is a measure of the significance of a peak by comparing the peak to its surroundings:

$$\lambda_i = r(w_0) - \max\left[r(b_{l,i}), r(b_{r,i})\right], \quad \begin{aligned} b_{l,i} &= \underset{j \in [0, L_0]}{\operatorname{argmin}} \left[r(w_j)\right], \\ b_{r,i} &= \underset{j \in [L_0, n-1]}{\operatorname{argmin}} \left[r(w_j)\right] \end{aligned} \tag{3}$$

where $r(w_0)$ is $-R^2$ at L_0 and b_l and b_r are the arrays of left and right bases of the peaks; we are making use of the notation by which $r(w) \equiv -\mathbf{R}^2$.

For each peak in $r(w)$, we calculate the prominence and discard the ones that are below a given threshold (λ_0). The particular value for this threshold was not critical for the overall performance of the algorithm. For the sake of stability, we use multiple window sizes (100, 150, 200 and 250 ms) and combine the results in the following way. For each window k, we have an array of candidate points ($\mathbf{x}^k_{cand}$). Then, we create a vector of weights ($\mathbf{w}_k \in \mathbb{R}^n$) that have a value equal to a Gaussian distribution centered around each candidate point of each window. Mathematically:

$$\mathbf{w}_k(x) = \begin{cases} \mathcal{N}(x - x^k_{cand}, \sigma) & \text{if} \quad x \in \mathcal{B}^k_{cand} \\ 0 & \text{otherwise} \end{cases}$$

where $\mathcal{B}^k_{cand}$ denotes the vicinity of each point in x_{cand} for window k. Finally, the RT is given by:

$$RT = \underset{x}{\operatorname{argmax}} \left[\left(\sum_k \mathbf{w}_k(x) \right) / x \right] \tag{4}$$

Fig. S1 shows the distribution of RTs split by trial outcome over sessions, for both species; Fig S2 shows example paths and detected RTs for both species.

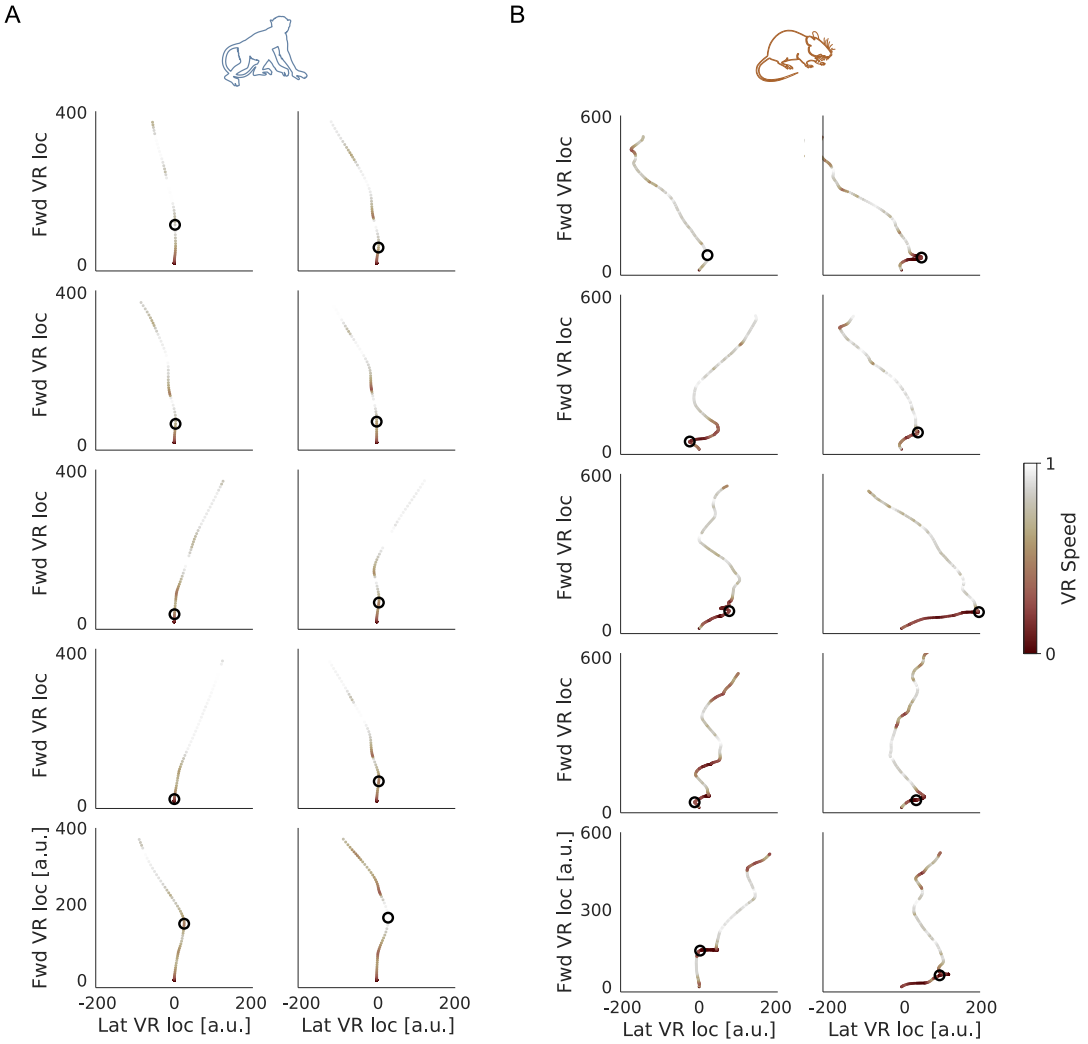

Fig. S2. Example VR paths Paths are colored according to the normalized running speed. **A)** Example paths for macaques, with the detected RT as circles. **B)** Same, but for mice.

Facial features. The extraction of the predictors for the MSLR model involves a multi-step process to go from continuous recording time (60 Hz for video data and 500 Hz for the macaque eye-tracker) to trial-based predictions.
First, we chose several points of interest on the animals' faces, which are then automatically identified and tracked over time using DeepLabCut (17):

- Macaque: both ears, eyebrows, nostrils and lips (see Fig. S3 A).

- Mouse: nose tip, left ear, left eye and median whiskers location (since we have a side view of the face, see Fig. S3 B).

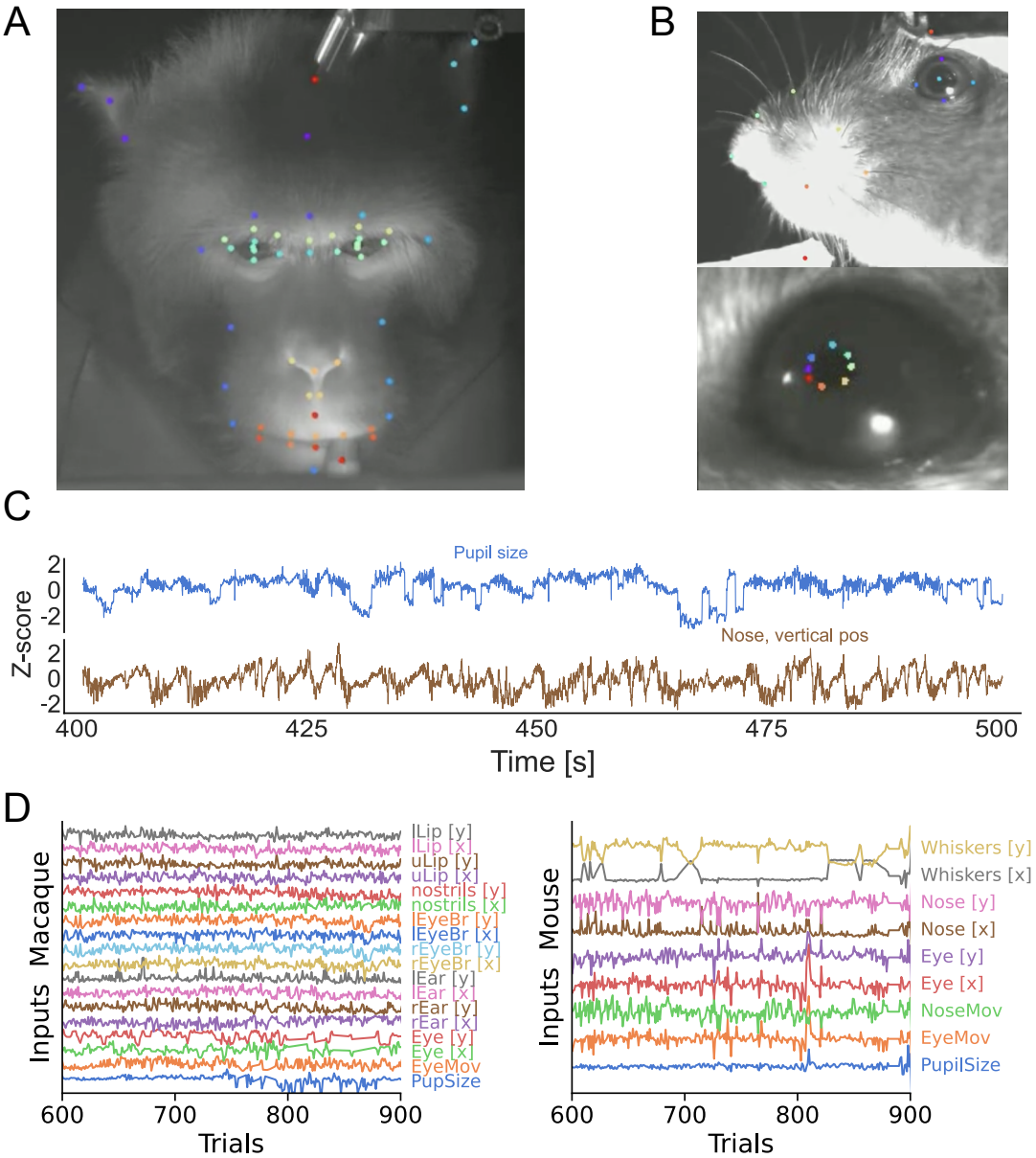

Fig. S3. Face features. A) Example frame of the macaque face camera. We have marked the key points that we used as the raw data for our pipeline. For this animal, we track a total of 73 key points. Some of them will be aggregated into centroids of interest, to minimize the influence of noise. **B)** Same as **A)**, but for the mouse. In this case, we also have a separate model for tracking pupil changes. **C)** Two example traces of a common feature for both species, over time. **D)** As described in the Methods, we use trial summaries for each of the face features of interest. Here, we show all of them, after having preprocessed them, for an arbitrary selection of 300 trials.

Once the data streams were aligned, we computed the median location (x, y) of each facial point over the 250 ms window before the stimuli appeared on the dome. This time window was chosen to make sure that all of the facial expressions of the animals are due to internally generated processing, rather than stimulus processing. Different window sizes (particularly: $200, 300$ and 500 ms) did not yield any qualitative difference. In addition to the median location, we also computed the total velocity of each facial point.

For both species, we further computed the median pupil size over the same time window. Pupil size is a well-known indicator of arousal and cognitive load, and thus provides valuable information about the internal state of the animal.

This resulted in a set of data points for each trial, corresponding to the median vertical and horizontal location, and total velocity of each of the facial features. These data points serve as the predictors for the MSLR model.

Markov-Switching Linear Regression. Markov-Switching Linear Regression (MSLR) models, which we ran using *Dynamax* (69), are a powerful tool for modeling time series data that exhibit regime-switching behaviour, where the underlying dynamics of the system change over time.

The MSLR model is defined by a set of linear regressions, each associated with a particular state of a discrete Markov chain. The state of the Markov chain determines which sets of weights and biases predicts the evolution of the observed data at each time step. The transitions between states are governed by the transition probabilities of the Markov chain, which are learned from the data.

Formally, an MSLR model can be described as follows. If S is the total number of latent (discrete) states of a Markov process, at each time step t, a given state $z_t\big(\in \{0,1,...,S\}\big)$ will follow a Markovian evolution such that:

$$z_{t+1}\big|z_t, \{\pi_s\}_{s=1}^{S} \sim \pi_{z_t} \tag{5}$$

where $\{\pi_s\}_{s=1}^{S}$ is the state transition matrix of the Markov process and π_s is the row corresponding to state s. As these are stochastic matrices, $\pi_s \in [0,1]^S$.

Let the $M-$dimensional input time series at time t be denoted by $x_t\big(\in \mathbb{R}^M\big)$. Let the $N-$dimensional output time series at time t be denoted by $y_t\big(\in \mathbb{R}^N\big)$. Then, in the case of a MSLR, the discrete latent variable at time t (z_t), will dictate which emission weights ($W \in \mathbb{R}^{N\times M}$) and emission biases ($b_s \in \mathbb{R}^N$) we will use to predict the outputs (emissions) based on the inputs (predictors). Moreover, an emission covariance matrix ($\Sigma_s \in \mathbb{R}^{N\times N}_{\succeq 0}$) will also have to be learnt. Explicitly, at time t, the emission distribution in this model is given by:

$$p(y_t \mid z_t, x_t, \theta) = \mathcal{N}(y_t \mid W_{z_t} x_t + b_{z_t}, \Sigma_{z_t}) \tag{6}$$

Therefore, the problem of fitting this model amounts to finding the set of emission parameters denoted by:

$$\theta = \Big\{\big(W_s, b_s, \pi_s, \Sigma_s\big)\Big\}_{s=1}^{S} \tag{7}$$

In other words, the aim is to find the weights (W_s) and biases (b_s) for the linear regressions and the transition π_s and covariance Σ_s matrices for the Markov process.

In our case, the discrete latent variable (z_t) represents the internal state of the animal at trial t, which is inferred from the facial features (x_t) extracted using DeepLabCut (17) and the observation (y_t) that represents the RT of the animal. We trained the MSLR model using the Expectation-Maximization (EM) algorithm (70), which iteratively estimates the state sequence and updates the model parameters to maximize the likelihood of the observed data. For further details, we refer the reader to (71). We iterated the EM algorithm for 500 times, for all models.

Training and inference. We used an $80:20$ ratio for train-test splitting and performed hyperparameter optimization by cross-validating the training set only (see *Model tuning* for details on CV and model selection). For each species, we concatenated the training sets of all sessions, with forced transitions in between the sessions (setting predictors and emissions to 0 for 50 consecutive trials), so that state probabilities are reset. Then, after optimizing each model, we performed inference on each held out test set (separately per session). We decided to take this approach for various reasons:

- Model generalization: as the model learns from potentially different faces, it is likely that it can pick up on common information between them.
- Model interpretability: given that we do not update the weights at the inference step, all internal states have the same meaning over subjects and, thus, are directly comparable.
- Better convergence: increasing the number of training samples (i.e. concatenating sessions as opposed to training a different model per session) allows the model to have more data to learn from.

All of the results in the main text, unless otherwise stated, are for held out data.

Model tuning. For the model we described in *Markov-Switching Linear Regression*, there are several parameters that can be tuned to explain the data better. In our case, we decided to explore the influence of changing the maximum number of internal states (S), to add sticky transitions to the Markov process (a self-bias term in the transition matrix π, making states taking longer to transition to a different one), and to vary the transition matrix sparsity (concentration).

In order to balance model performance with scientific insights, we took a hybrid approach. We increased the number of internal states in a greedy way, to show that the error saturates and that there are diminishing returns when increasing model complexity. On the other hand, for a given number of states, we optimized two free parameters of the Markov process: state stickiness and state concentration. For the sake of efficiency, we used Optuna (72), a flexible framework to implement Bayesian optimization. In Table 1 we report the relevant quantities for this process.

To select the best combination of parameters, we performed 5-fold Time-Blocked Cross-Validation (73).

Parameter	Value
Concentration	$[0, 100]$
Stickiness	$[0, 100]$
Sampler	CMA-ES (74)
Objective function	R^2
Number of Searches	100

Table 1. Parameter values for the Bayesian parameter optimization procedure. These are independently explored for each number of internal states of the HMM.

Synthetic data and ground truth states. In order to validate the retrieval of states when we do not have access to ground truth ones, we generated a time series of ground truth emissions and states based on the given inputs (by using the same input data as in the main text). To this end, we trained an MSLR model with a known given number of states and sampled some emissions and states sequence from it. We aimed to recover the appropriate number of states with the correct temporal sequence, and to correctly predict the emissions. Figure S4A illustrates the input data (composed of session-concatenated mouse facial features, as described in the main text). In Figure S4C, we show that, once we have selected the appropriate number of states, the model's log-probability does peak at the ground-truth one (dashed vertical line). In Figure S4D, we show a comparison between the true and the inferred states, for some example trials. Although the temporal coincidence of the state transitions is very high, due to the stochastic nature of the model, some state labels might be permuted (i.e., state 1 in our model might correspond to state 0 in the ground truth states). Therefore, in order to quantify state similarities and to account for state-swapping, we one-hot encoded the true and predicted states sequences and correlated all pairs with each other (Figure S4D). There is an almost perfect match ($\rho(s_{true}, s_{pred}) > 0.9$) between the true and inferred states (the 99th percentile of the surrogate correlation distribution was 0.12).

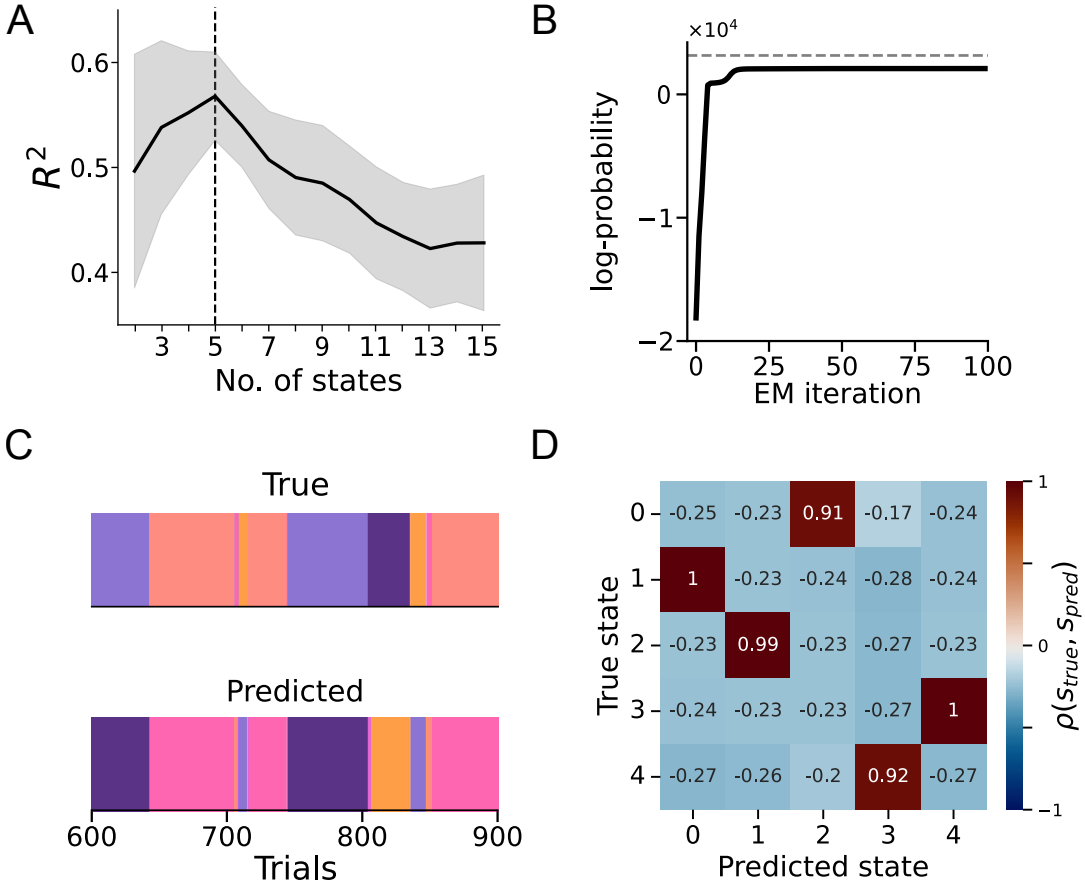

Fig. S4. Synthetic states and emissions. A) Performance when varying the number of states. We are able to recover the number of states (vertical dashed line) that generated the ground truth emissions. **B)** For the selected number of states, log-probability of the fitted parameters; it converges to the ground truth value (horizontal dashed line). **C)** Some example trials for the true and predicted states. State transitions are correctly captured, but state labels might be permuted. **D)** Temporal correlation between the one-hot encoded state arrays. There is an almost perfect match in between the predicted and the true states arrays, up to a label permutation.

ARHMM. As we wanted to ensure that facial features were indeed informative of reaction times (RTs) beyond what is to be expected by the RT autocorrelation structure, we implemented an Auto-Regressive Hidden Markov Model (ARHMM). In this

case, we used the same pipeline as we detailed in the previous sections, but substituted the facial features at the current trial t for the RT of the previous trial $(t-1)$. As it can be seen from Fig. S5, the facial features model outperforms the ARHMM for all states, for both species. The ARHMM also generalizes worse than the MSLR face model (inset, test R^2).

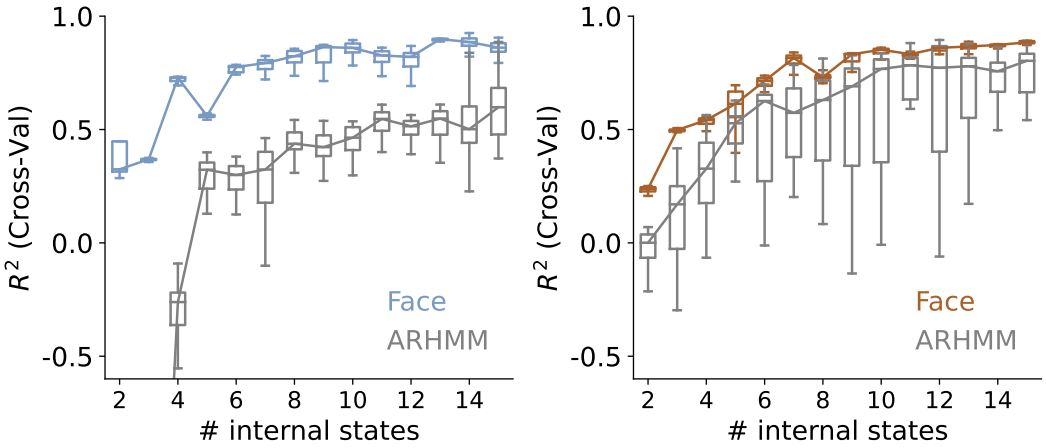

Fig. S5. Comparison of the MSLR face model and the reaction time Auto-Regressive HMM. Both face feature models outperform their Autoregressive counterparts, for any number of internal states that we swept over. Nevertheless, it can be seen that the performance gap is smaller in mice than in macaques. This is consistent with the finding that mice are more history dependent than macaques (See Fig. 5 D).

Task performance and internal states. We were interested in investigating whether the inferred internal states were correlated with task performance, even though the model had not been trained on such information. We therefore used the predicted state probabilities to decode choice, using a simple Logistic Regression model, with a $L2$ penalty term. After verifying that the model does indeed classify outcome beyond chance level (Fig. S6), we took the weight of each state as a proxy for how related it was to each outcome.

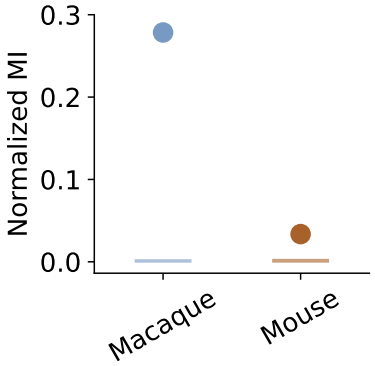

Fig. S6. Inferred state probabilities decode outcome beyond chance. We used a normalized version of Mutual Information that already takes chance level into account and sets that as 0.

Supplementary figures

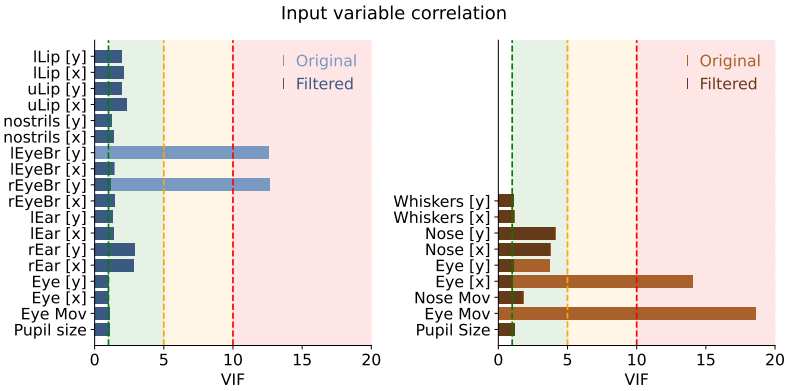

Fig. S7. Input variable correlation. We show that, out of all of the original variables (in lighter colors), we end up discarding one per animal (*Left Eyebrow [y], macaques; Eye movement, mice*), given that they were highly multi-colinear with some of the other predictors, as measured by the Variance Inflation Factor (VIF). After discarding them and recomputing the VIF, we did not find any alarming colinearity.

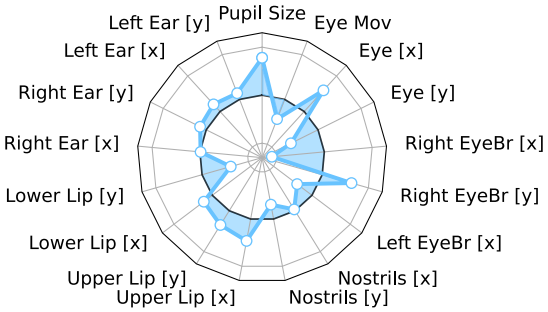

Fig. S8. Face features importance for State B in macaques.

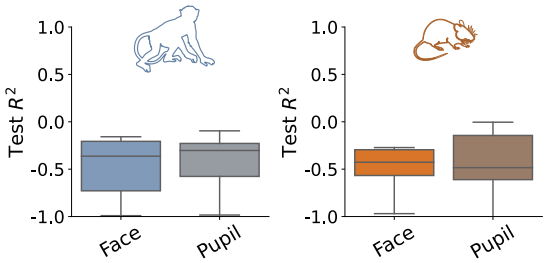

Fig. S9. No-switching model performances. If we assume there is a simple linear relationship (i.e. we do not allow for any switching) between the face (or the pupil) and the Reaction Time, the test performance is back at chance level.

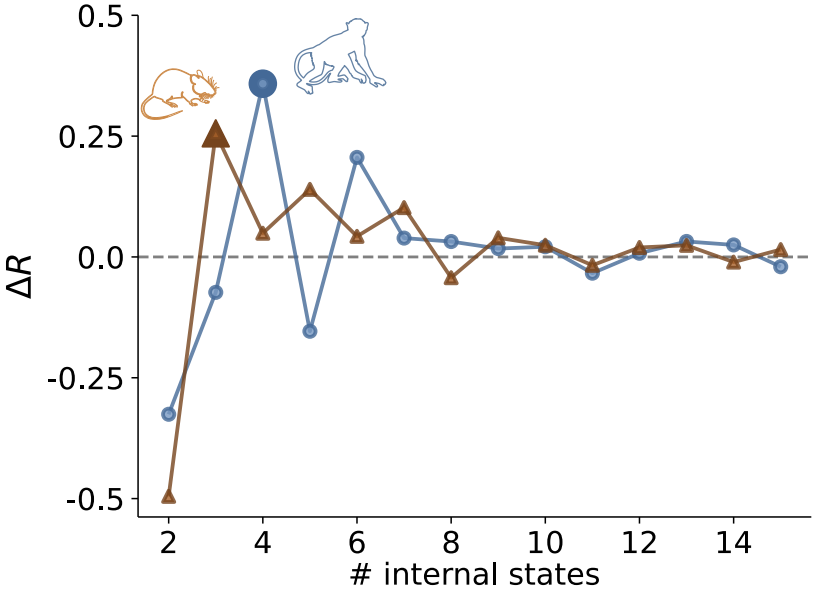

Fig. S10. Finite difference performance curves. We selected the number of states that was given by the maximum relative gain in performance, computed as the finite difference in Cross-Validated R^2.

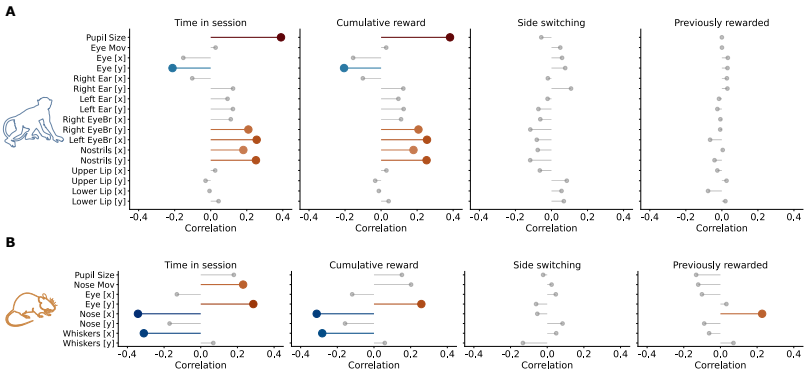

Fig. S11. Facial features and task history correlations, for macaques (top) and mice (bottom): time spent in session, cumulative reward in the session, the correct stimulus switching from being the left to the right one compared to the previous trial and vice versa, and whether the previous trial had been rewarded or not. Those highlighted in colour are significantly correlated.

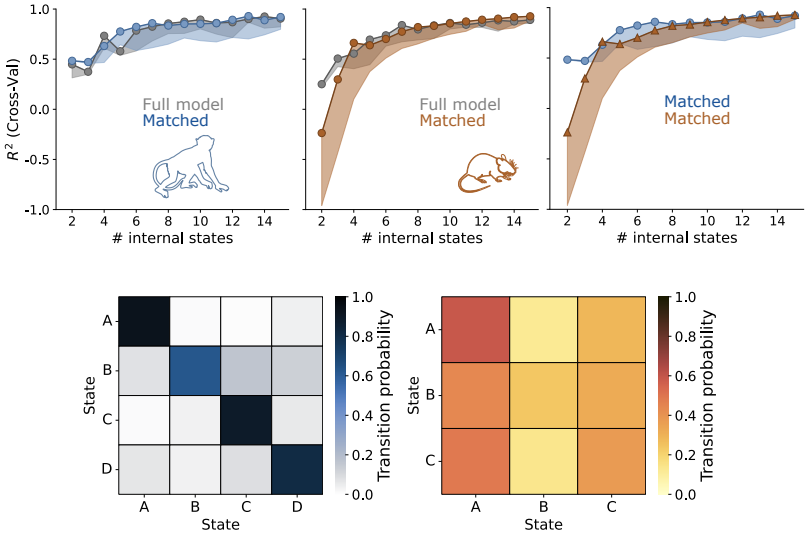

Fig. S12. Matched models. Our main findings still hold if we match the number of predictors, trials and animals we use in each species.

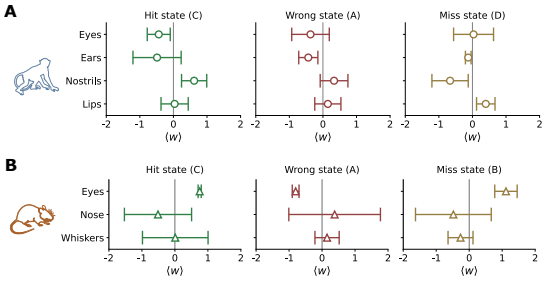

Fig. S13. Simplified face features weights.

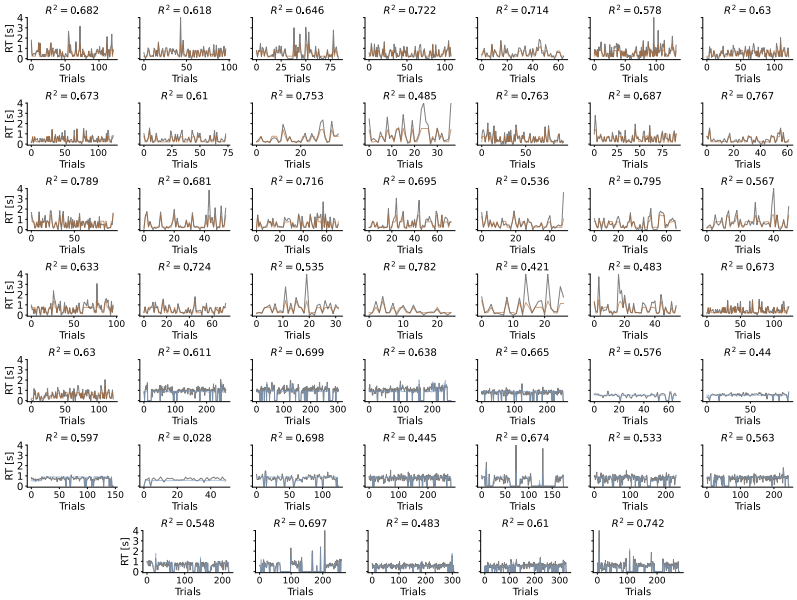

Fig. S14. Predicted and true RTs for all sessions, for the held out test set. For both species, true RTs are shown in gray; for mice, predicted RTs are shown in orange; for macaques, predicted RTs are shown in blue. Each subplot is titled with that session's model performance (R^2).

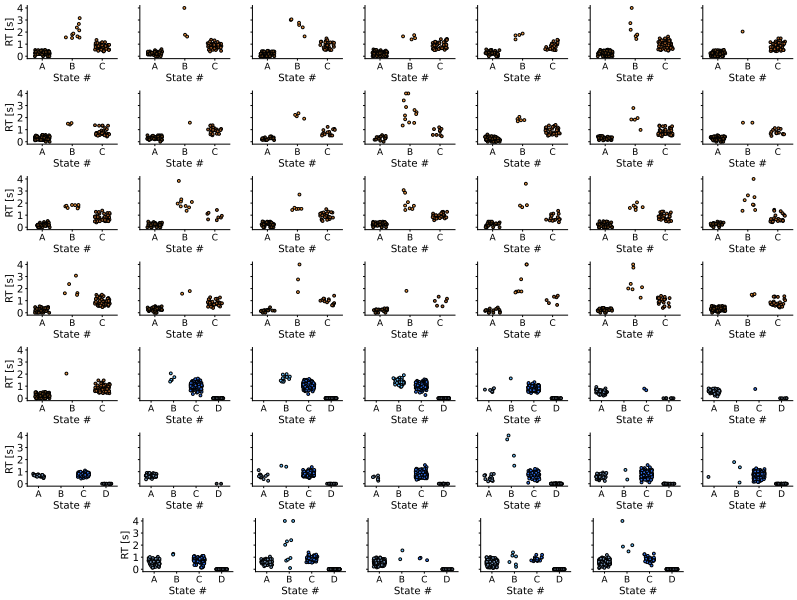

Fig. S15. RTs over states for all sessions, for the held out test set. For mice, distributions are shown in orange; for macaques, distributions are shown in green-blue.

Fig. S16. Most likely states for all sessions, for the held out test set. For mice, states are shown in orange; for macaques, states are shown in green-blue.