Article

Sensory expectations shape neural population dynamics in motor circuits

https://doi.org/10.1038/s41586-025-09690-9

Received: 24 January 2025

Accepted: 29 September 2025

Published online: 29 October 2025

Check for updates

Jonathan A. Michaels[1,2,3]✉, Mehrdad Kashefi[1], Jack Zheng[1,2], Olivier Codol[1], Jeffrey Weiler[1], Rhonda Kersten[2], Jonathan C. Lau[1,4], Paul L. Gribble[1,2,5], Jörn Diedrichsen[1,6,7] & J. Andrew Pruszynski[1,2,5]✉

The neural basis of movement preparation has been extensively studied during self-initiated actions, in which motor cortical activity during preparation shows a lawful relationship to the parameters of the subsequent action[1,2]. However, movements are regularly triggered or corrected on the basis of sensory inputs caused by disturbances to the body. Since such disturbances are often predictable, and since preparing for disturbances would make movements more precise, we hypothesized that expectations about sensory inputs also influence preparatory activity in motor circuits. Here we show that when humans or monkeys are probabilistically cued about the direction of future mechanical perturbations, they incorporate sensory expectations into their movement preparation and improve their corrective responses. Using high-density neural recordings, we establish that sensory expectations are widespread across the brain, including the motor cortical areas involved in preparing self-initiated actions. The geometry of these preparatory signals in the neural population state is simple, and scales directly with the probability of each perturbation direction. After perturbation onset, a condition-independent signal shifts the neural state leading to rapid responses that initially reflect sensory expectations. Using neural networks coupled to a biomechanical model of the arm[3], we show that this neural geometry emerges only when sensory inputs signal that a perturbation has occurred, before resolving the direction of the perturbation. Thus, just as preparatory activity sets the stage for self-initiated movement, it also configures motor circuits to respond efficiently to sensory inputs.

Humans and animals are often able to prepare a movement in advance and such preparation generally makes movements more precise. The neural basis of movement preparation and its relationship to movement execution has frequently been studied with delayed action paradigms, in which the nature of a future movement is instructed but its execution must wait until a subsequent go cue (reviewed in refs. 1,2). During the preparatory period, between the movement instruction and the go cue, muscle activity remains unchanged but motor cortical activity represents parameters of the future movement[4–11], predicts movement variability[12] and reaction time[13–16], and is causally linked to motor execution[17,18], presumably by setting the initial state of the dynamical system that ultimately produces movement[19–21].

Although preparing specific movement parameters is an essential aspect of self-initiated actions, movements are regularly triggered or corrected on the basis of sensory inputs caused by disturbances to the body or environment. Since such disturbances can often be predicted, and since preparing for potential disturbances would improve motor performance, we hypothesized that sensory expectations should also directly shape preparatory activity in motor cortical circuits. Such a scheme is a key prediction of theories of biological motor control based on optimal feedback control[1,22,23] and would be consistent with previous reports that motor cortical areas rapidly respond to sensory inputs (reviewed in refs. 24,25) in a way that accounts for biomechanical[26] and task constraints[27–32].

Here we show that when cued about the likely direction of future mechanical perturbations, humans and macaque monkeys readily incorporate expectations about the upcoming sensory input into their movement preparation and that this preparation improves their performance. We then demonstrate that information about sensory expectations is robust and widespread in monkey motor circuits, but not early sensory areas. The neural geometry of these signals is simple, directly representing the probability of each perturbation direction. Finally, we develop a normative model of the motor system that shows how this neural geometry is beneficial for countering perturbations and how it relies on the timing of incoming sensory signals.

[1]Western Centre for Brain and Mind, Western University, London, Ontario, Canada. [2]Department of Physiology and Pharmacology, Western University, London, Ontario, Canada. [3]School of Kinesiology and Health Science, York University, Toronto, Ontario, Canada. [4]Department of Clinical Neurological Sciences, London Health Sciences Centre, Western University, London, Ontario, Canada. [5]Department of Psychology, Western University, London, Ontario, Canada. [6]Department for Computer Science, Western University, London, Ontario, Canada. [7]Department for Statistical and Actuarial Sciences, Western University, London, Ontario, Canada. ✉e-mail: jmichae@yorku.ca; andrew.pruszynski@uwo.ca

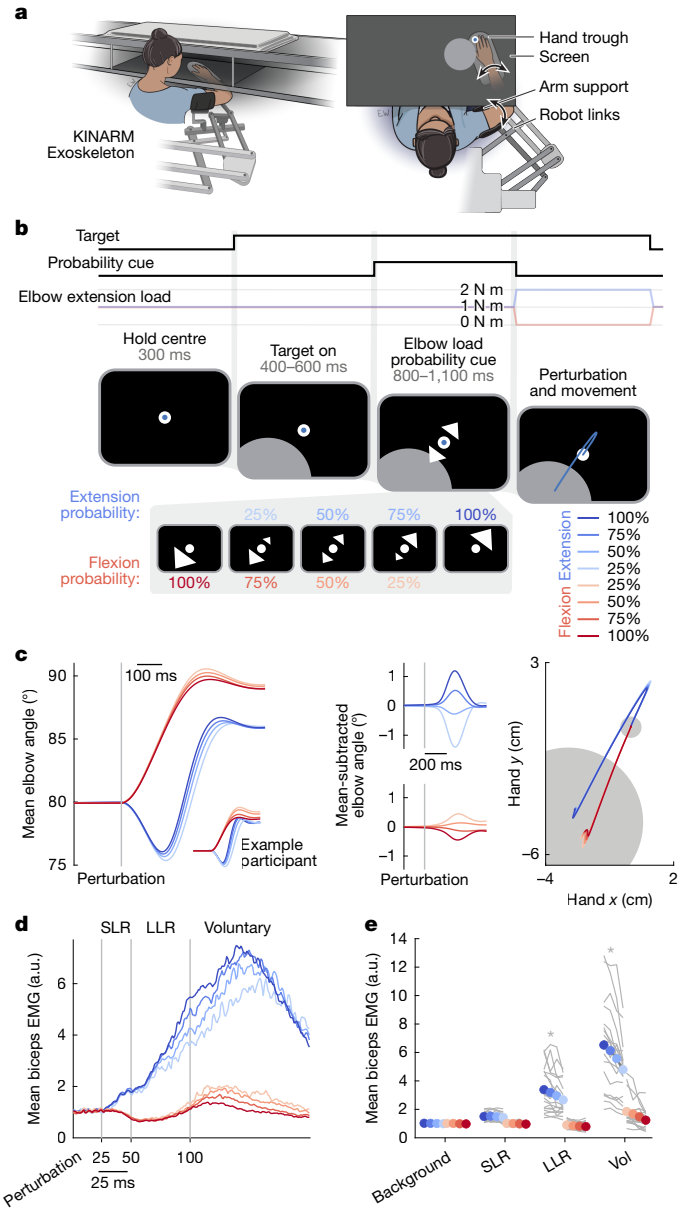


Fig. 1 | Long-latency stretch reflexes are sensitive to sensory expectations. **a**, Human participants were seated in the KINARM exoskeleton robot, allowing presentation of visual stimuli, tracking of the arm and application of mechanical forces to the shoulder and elbow. Image credit: Erika Woodrum, Erika Woodrum Art, Ltd. **b**, Participants held their hand at a small central target while resisting a background load and were shown a peripheral goal target. On single trials participants received one of five visual cues indicating the probability that the upcoming mechanical perturbation would flex or extend their elbow joint and thus push their hand into or out of the goal target. After a variable delay, a perturbation drawn from the displayed probability distribution was applied and participants had to respond to the perturbation by moving their hand into the goal target quickly and accurately. Note that the five visual cues ultimately map onto eight experimental conditions because three of the visual cues (that is, 25%/75%, 50%/50% and 75%/25%) lead to two possible perturbations. **c**, Mean elbow kinematics across participants ($n = 20$) showed that participants responded to the perturbations in a graded manner, in which the speed with which participants moved to the target depended on the cued probability of each perturbation. Middle, the mean-subtracted elbow kinematics. Right, the mean hand kinematics. **d**, Mean biceps muscle activity measured via surface electromyography (EMG) as a function of perturbation direction and probability cue. Note that the initial excitatory response to extension perturbations in the short-latency reflex (SLR) window is the same for all probability cues, but that the subsequent excitatory response in the long-latency reflex (LLR) and voluntary windows is scaled by the probability cues. a.u., arbitrary units. **e**, Mean muscle responses were significantly modulated by probability cues in the long-latency reflex and voluntary windows (repeated-measures ANOVA, $P < 0.001$).

Reflexes reflect sensory expectations

To investigate whether expected sensory inputs shape preparatory activity in motor cortical areas, we designed a task in which human participants were given probabilistic information about how a future mechanical perturbation would displace their arm. Participants ($n = 20$) sat in an exoskeleton robot and maintained the position of their hand within a small central target while countering a constant background load (Fig. 1a and Methods). On each trial, they were randomly shown one of five visual cues that explicitly informed them about the probability that their elbow joint would be flexed or extended by an upcoming mechanical perturbation (Fig. 1b). After an unpredictable delay, a perturbation drawn from the cued probability distribution was applied to their elbow joint. Participants were instructed to respond to the perturbation by moving their hand into a large goal target as quickly and accurately as possible.

Participants were very good at the task (success rate: 85.7 ± 6.3%; mean ± s.d.), and rarely initiated a movement incorrectly during catch trials without perturbations (error rate: 5.0 ± 3.5%). The probability cues biased task performance in a graded manner, with participants countering perturbations most effectively if the visual cue indicated with 100% certainty that the particular perturbation direction would occur, with a graded response for lower-probability cues (Fig. 1c, left). This effect was even more visible in the mean-subtracted elbow kinematics (Fig. 1c, middle). That is, elbow displacement decreased for the same perturbation as the probability of that perturbation increased. Statistical analysis showed a reliable difference in elbow kinematics as a function of probability in the movement period (0–400 ms after perturbation onset; repeated-measures ANOVA, $F(3,19) = 59.75$, $P < 10^{-6}$).

It is unlikely that participants had been making binary guesses in line with the most likely perturbation on individual trials, causing the average to show a graded response, as the distribution of elbow velocity inversion times following perturbation onset generally tended towards slower responses as perturbations became less likely, showing unimodal distributions (see Extended Data Fig. 1 for individual participant distributions). Indeed, these distributions did not significantly differ from a unimodal distribution for any participant or probability (Hartigan's dip test on extension conditions, all $P > 0.13$), although we may not have the statistical power to fully exclude this possibility in all participants.

We specifically chose mechanical perturbations that evoked stretch reflexes because the various components of the stretch reflex involve different neural circuits[33]. The short-latency component, measured as muscle activity occurring 20–50 ms after muscle stretch, is generated entirely by spinal circuits[33]. The long-latency component, measured as muscle activity occurring 50–100 ms after muscle stretch, includes a contribution from motor cortical areas via the transcortical feedback pathway[24,34]. Thus, if motor cortical circuits are set in accordance with sensory expectations, long-latency stretch reflexes should be sensitive to the probabilistic information about the direction of the future mechanical perturbation. Indeed, previous studies using a similar approach have shown that the long-latency stretch reflex is shaped by many parameters that influence preparatory activity in motor cortical areas in the context of self-initiated movements[26–29].

As expected, our mechanical perturbations evoked excitatory or inhibitory stretch reflex responses starting around 20 ms after muscle stretch, depending on whether the perturbation lengthened or shortened the muscle (Fig. 1d and Extended Data Fig. 2). Notably, stretch reflex responses were initially insensitive to probability information, which emerged in a graded manner approximately 70 ms after muscle stretch, within the long-latency epoch associated with the transcortical

feedback pathway (Fig. 1e). We found no reliable effect of probability in the short-latency epoch (short-latency reflex, 20–50 ms after perturbation onset; repeated-measures ANOVA, $F(3,19) = 2.35$, $P = 0.08$). By contrast, we did find a reliable effect of probability in both the long-latency (long-latency reflex, 50–100 ms; $F(3,19) = 7.55$, $P = 0.0002$) and voluntary (100–150 ms; $F(3,19) = 18.28$, $P < 0.0001$) epochs. Note that this modulation did not reflect anticipatory modulation of muscle activity prior to the perturbation[35], as there was no reliable effect of probability on muscle activity immediately prior to muscle stretch (background, −200 to 0 ms; $F(3,19) = 0.86$, $P = 0.46$).

Together, these results show that participants prepare perturbation responses on the basis of sensory expectations and suggest that this process engages motor circuits via the transcortical feedback pathway.

Sensory expectations shape neural activity

Our finding that long-latency stretch reflexes are sensitive to sensory expectations strongly suggests motor cortical involvement but cannot establish a direct link to its preparatory state. To directly test whether sensory expectations shape preparatory activity in motor cortical areas, we trained two macaque monkeys in a similar version of the task performed by human participants (Fig. 2a and Methods). In this version, monkeys returned their hand to a single central target when perturbations flexed or extended their elbow joint. Unlike human participants, who were verbally informed about the probabilistic nature of the cues and enacted the association immediately, monkeys learned to associate the visual cues with probabilities through experience over several months.

After training, monkeys were very good at the task, completing a large percentage of initiated trials correctly (monkey M: 89.1 ± 4.1%, monkey P: 93.2 ± 1.9%; mean ± s.d. over sessions) and rarely starting their response before perturbation onset (monkey M: 7.3 ± 3.1%, monkey P: 5.4 ± 1.6%). Similar to human participants, monkeys showed a clear effect of the probability cue on their elbow kinematics in the 400 ms following the perturbation (Fig. 2b; one-way ANOVA of single-trial elbow angle over probability conditions, monkey M: extension $F(3,16424) = 79.6$, $P < 10^{-6}$, flexion $F(3,16274) = 169.3$, $P < 10^{-6}$; monkey P: extension $F(3,7099) = 61.8$, $P < 10^{-6}$, flexion $F(3,7054) = 35.5$, $P < 10^{-6}$), demonstrating that they learned to prepare for future perturbations on the basis of the probability cue. Also similar to human participants, it is unlikely that these differences were caused by guessing what perturbation would occur on individual trials, since the distribution of elbow velocity inversion times following elbow perturbations did not significantly differ from a unimodal distribution (Hartigan's dip test, trials pooled across sessions and each condition tested separately; monkey M: all P values > 0.81; monkey P: all P values > 0.64; see Extended Data Fig. 3 for individual distributions).

Having established that monkeys use probabilistic cues to prepare their responses to mechanical perturbations, we recorded single neurons from multiple brain areas to assess how neural population activity supported this preparation. To do so, we developed a new recording setup allowing parallel single neuron extracellular recording using Neuropixels probes (Methods). In both monkeys, we recorded from four areas potentially involved in this preparation (Fig. 2c): primary somatosensory cortex (S1), primary motor cortex (M1), dorsal premotor cortex (PMd) and dorsolateral prefrontal cortex (dlPFC), recording from 8,141 single neurons in total (see Extended Data Table 1 for a breakdown of all recording sessions).

We used demixed principal component analysis (dPCA) to decompose neural population activity into task-related components, including probability information (that is, the visual cues), perturbation direction (elbow flexion or extension), condition-independent changes, and any linear interactions. To fit dPCA, neurons within each brain region were pooled across recording sessions. Once these dimensions were found, they were used to project simultaneously recorded neurons onto dPCs for single trials. Figure 2d shows the single-trial correlation between the top probability dimension of each area and the perturbation probability throughout the preparation time. Between the cue and perturbation time, the top probability dimension was predictive of probability to some extent in each area, however, there were clear differences across areas. Notably, PMd and dlPFC showed a single-trial correlation with probability around 100 ms after the cue was presented, followed by ramping probability information in M1. By contrast, representations in S1 were very small and ramped up latest during preparation.

To visualize the nature and extent of these representations, we plotted the condition-averaged neural population activity of each area projected into the top dPC of each factor (Fig. 2e,f). During preparation, conditions corresponding to identical visual cues overlap in the probability dimension, leading to five distinct clusters of traces. These probability representations were clearest in PMd and M1, reaching their peak shortly before the perturbation and then decreasing throughout the movement period. As expected, the probability dimensions represented only probability information and did not delineate what perturbation was actually applied, demonstrating that probability information was linearly separable from perturbation information. We next considered how the neural geometry was structured. Although dPCA did not look for any particular ordering of the cues, the top probability dimension clearly ordered these cues by the probability magnitude, especially in M1 and PMd. The top perturbation dimension clearly separated the two perturbation directions in all areas except in dlPFC, where representation of perturbation direction was small. A supplementary analysis confirmed that perturbation direction information was evident earliest in S1, followed by M1 and PMd (Extended Data Fig. 4a).

Also notable is that all areas showed dominant condition-independent signals, which vary across time but not in relation to other task variables. These varied across areas, with S1 and M1 showing the sharpest responses to the perturbation, while higher order areas (PMd and dlPFC) showed less sharp responses to the perturbation and greater ramping during preparation. The origin of the condition-independent signal at perturbation onset is unclear (see Discussion), since the majority of sensory inputs from the periphery would be directional and show an opposite response for different elbow perturbation directions.

To summarize these results across all dPCs, we examined the total variance explained by each factor (including interactions and residual variance) across areas and monkeys (Fig. 2g,i). We found a clear gradient of probability information across areas, with probability information least dominant in S1 and becoming more dominant in M1, PMd and dlPFC. Conversely, we found the reverse gradient for perturbation information, from most dominant in S1 to least dominant in dlPFC. The interaction of probability and perturbation explained only a very small amount of variance overall. Notably, condition-independent dimensions represented the majority of variance explained in all areas, highlighting the importance of these time-varying signals and in line with previous work during self-initiated movements[36,37]. A supplementary analysis showed that condition-independent perturbation responses were generally evident earlier in S1 than in M1 and PMd (Extended Data Fig. 4b). Residual variance not captured by dPCA was very small, with the exception of dlPFC, suggesting that the remaining variance is either not linearly separable or related to other factors not experimentally controlled.

Finally, if the probability dimension causally relates to the behavioural response to the perturbation, we would expect that fluctuations in the probability dimension on single trials would affect behaviour on the same trials. To test this idea, we used the single-trial projections of neural activity onto the first probability dimension and correlated them with elbow velocity in the first 150 ms following the perturbation period (Fig. 2h,j). In previous analyses, some probability information was present in dlPFC and S1. However, we did not observe any clear correlation between single-trial fluctuations in probability information

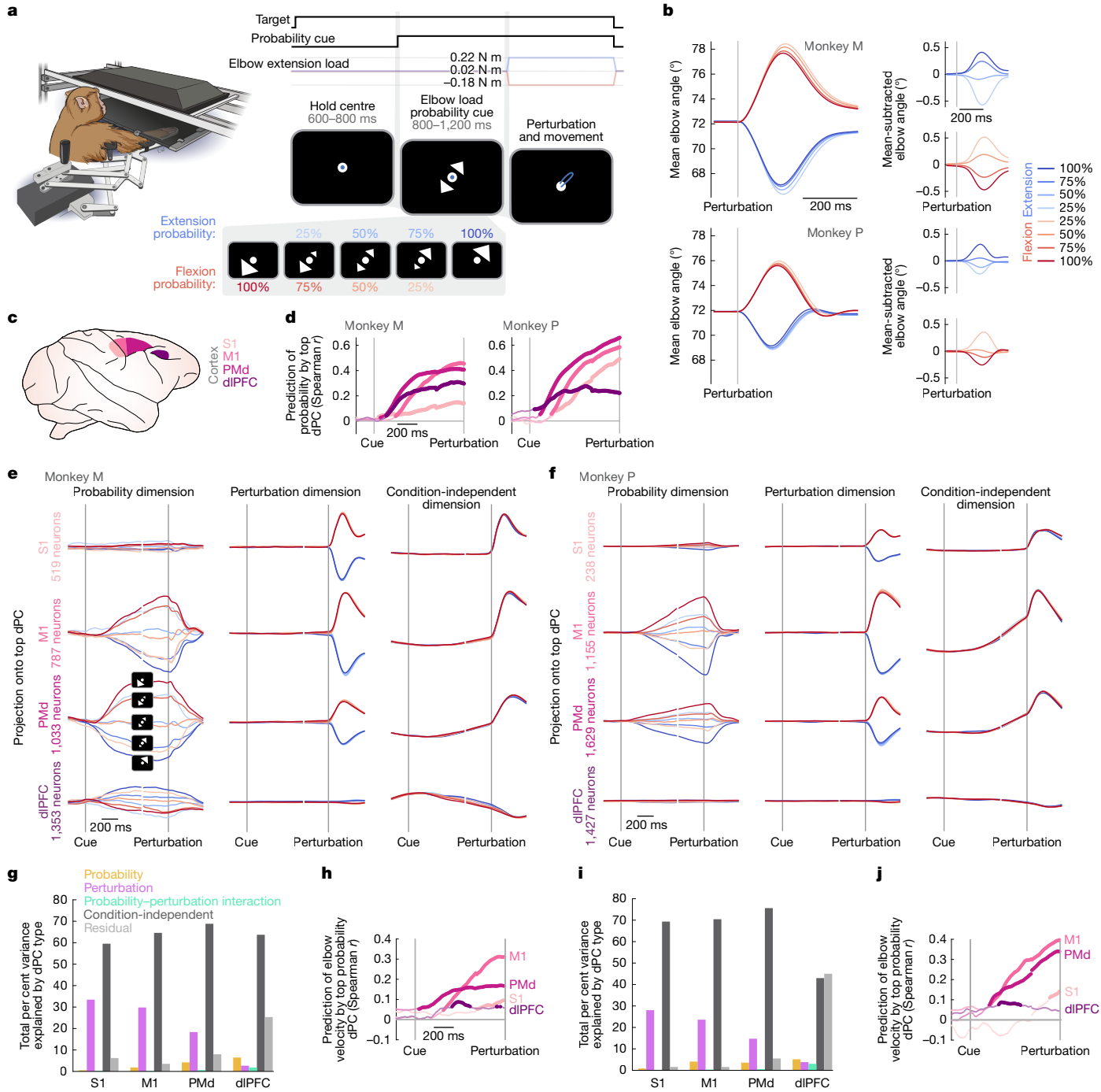


Fig. 2 | Sensory expectations shape preparatory activity in motor cortical areas. a, Two macaque monkeys performed a version of the task similar to the task performed by human participants, using a non-human primate version of the KINARM exoskeleton robot, needing to return to a central target after having their arm displaced out of the target by an elbow perturbation. Image credit: Erika Woodrum, Erika Woodrum Art, Ltd. **b**, Mean elbow kinematics across recording sessions showed clear scaling by probability cue in both perturbation directions. **c**, Single neurons were recorded over sessions from four cortical areas: S1, M1, PMd and dlPFC. **d**, dPCA was performed across the single neurons of each area to disentangle task-related signals at the population level. Projections onto the top probability demixed principal component (dPC) were used to predict probability magnitudes on single trials pooled across all sessions of each area. Thick coloured lines indicate correlations significantly greater than chance, where chance level was obtained by randomly shuffling probability conditions and correlating with the true probability condition (one-sided permutation test, $P < 0.001$, 10,000 iterations). **e**,**f**, Top dPCs are plotted for each task factor (probability, perturbation, condition-independent) across areas for monkey M (**e**) and monkey P (**f**). dPCs are normalized to the maximum value of each task factor pooled across areas. **g**, Summary of total variance explained per task factor in each area for monkey M. **h**, Mean prediction of single-trial elbow velocity in the 150 ms after the perturbation by the top probability dPC, performed separately for each perturbation direction and averaged, chance level calculation as in **d**. **i**,**j**, The same analyses as in **g** (**i**) and **h** (**j**), for monkey P.

and subsequent elbow velocity after the perturbation. By contrast, neural activity immediately before the perturbation in both M1 and PMd significantly predicted elbow velocity after the perturbation on single trials, suggesting that the probability representation in these areas may directly modulate rapid responses to perturbations, a result consistent with their key role in the transcortical feedback pathway.

Given the widespread presence of sensory expectation signals in multiple cortical areas, we extended our recording setup to enable

the use of 4.5 cm Neuropixels[38] in monkey P, to target subcortical and cortical medial wall structures. Thalamus is particularly interesting, given the rich and differentiated input and output connectivity of its various nuclei. We targeted the ventroposterior lateral thalamus (VPL), the ventral lateral posterodorsal thalamus (VLpd), and the ventral lateral anterior thalamus (VLa). VPL receives tactile and proprioceptive information from the cuneate nucleus and has dense projections to S1 (ref. 39) and M1 (ref. 40), whereas VLpd receives primarily cerebellar input and projects to the the supplementary motor area (SMA) and PMd/M1 (ref. 41), and VLa receives pallidal input and projects to SMA/PMd/M1 (refs. 41,42), motivating us to also record from SMA. Performing the same dPCA analysis as in Fig. 2, we found that probability information was present in most of these areas rapidly after cue presentation (Fig. 3b), reinforcing that this information is widespread in the motor circuit. Specifically, VLa and VLpd showed an increase in probability prediction in the top probability dPC within 100 ms of the cue, followed closely by SMA. By contrast, the VPL showed only low levels of prediction much later in preparation, similar to what we observed in S1. Note that observing low levels of probability decoding later during preparation in VPL relative to S1 does not contradict the expected hierarchy of these areas, since any preparatory-related probability activity is likely to be a consequence of top-down signals.

A summary of variance explained (Fig. 3d) shows that, as in previous analyses, condition-independent information was by far the most dominant. SMA and VPL had the largest perturbation responses, and probability information in VPL accounted for less than 0.3% of total variance. A supplementary analysis of perturbation responses showed that the timing of perturbation information in VPL was similar to S1 (Extended Data Fig. 4). Visualizing the projections in the top dPC for probability (Fig. 3c) confirms that probability information was minimal in the VPL. It is noteworthy that the geometry of probability information in the other thalamic nuclei and SMA were organized by relative probability, as observed in other cortical areas (Fig. 2). Post-perturbation elbow velocity was weakly predictable from the top probability dPC of all these areas close to perturbation time (Fig. 3e), suggesting some direct dependence between neural state and subsequent behaviour, although this relationship was much weaker than what was observed in PMd and M1.

Together, these results point to a strong and widespread representation of sensory expectations in cortical and subcortical motor circuits but not early sensory areas.

Sensory expectations accumulated from experience

In our previous experiments, humans and monkeys were shown a random probability cue on each trial, making the visual cue the only source of information for forming sensory expectations. Although examples like this exist in natural environments, expectations about future sensory inputs can also be discovered through experience interacting with the environment. Therefore, we designed an additional experiment to investigate the neural representation of sensory expectations acquired over multiple trials. In this experiment (Fig. 4a and Methods), monkeys received alternating blocks of randomly cued probabilities (presented visually) and adaptation blocks where perturbations were drawn from a single randomly chosen probability distribution that was not visually cued to the monkey (for example, 75% extension). We did not include the 100% probability conditions in the adaptation blocks to maintain some uncertainty about the underlying distribution. In adaptation blocks, the only way to determine the underlying perturbation probability distribution was by experiencing a succession of perturbation trials. In monkey M, we recorded 333 neurons in S1, 399 in M1, 244 in PMd, and 1,329 in dlPFC. In monkey P, we recorded 215 neurons in S1, 586 in M1, 1,629 in PMd, 1,427 in dlPFC, 131 in VPL, 204 in VLpd, 497 in VLa, and 413 in SMA.

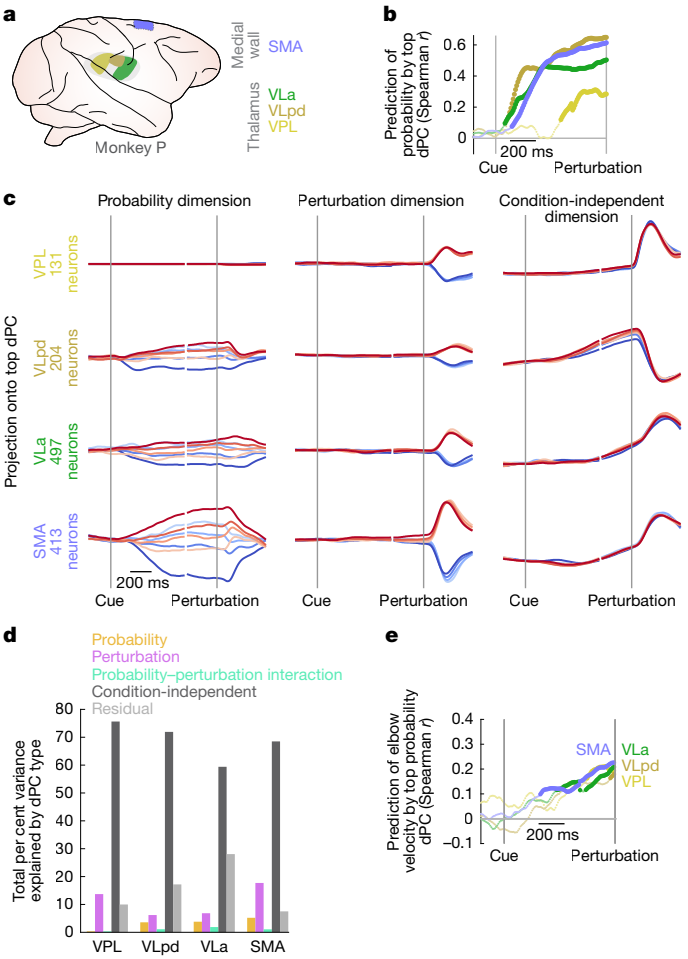


Fig. 3 | Sensory expectations shape preparatory activity in motor but not sensory thalamus. a, Single neuron recordings were made from monkey P over sessions from four areas: SMA in the medial wall, and the VLa, VLpd, and VPL nuclei of the thalamus. **b**, dPCA was performed across the single neurons of each area to disentangle task-related signals at the population level. Projections onto the top probability dPC were used to predict probability magnitudes on single trials pooled across all sessions of each area. Thick coloured lines represent significant correlations, where chance level obtained by randomly shuffling probability conditions and correlating with the true probability condition (one-sided permutation test, $P < 0.001$, 10,000 iterations). **c**, Top dPCs are plotted for each task factor (probability, perturbation, condition-independent) across areas. dPCs are normalized to the maximum value of each task factor pooled across areas. **d**, Summary of total variance explained per task factor in each area for monkey P. **e**, Mean prediction of single-trial elbow velocity in the 150 ms after the perturbation by the top probability dPC, performed separately for each perturbation direction and averaged, chance level calculation as in **b**.

Both monkeys adapted to the distribution of experienced perturbations, scaling their responses on the basis of the probability of the underlying distribution (Fig. 4b,c). To dissect the time course of this adaptation, we aggregated all of the visually cued trials across sessions (for each monkey and perturbation direction separately) and fit linear regressions to predict the probability distribution of each trial on the basis of the post-perturbation kinematics (shoulder and elbow velocity 0–400 ms post-perturbation; L1 regularization coefficient selected using Bayesian optimization with 5-fold cross-validation). We used this regression fit on the visually cued blocks to predict probability conditions on adaptation trials and assessed performance as the Spearman correlation for trials at all points within an adaptation block (Fig. 4d,e). For both monkeys, post-perturbation kinematics started showing a

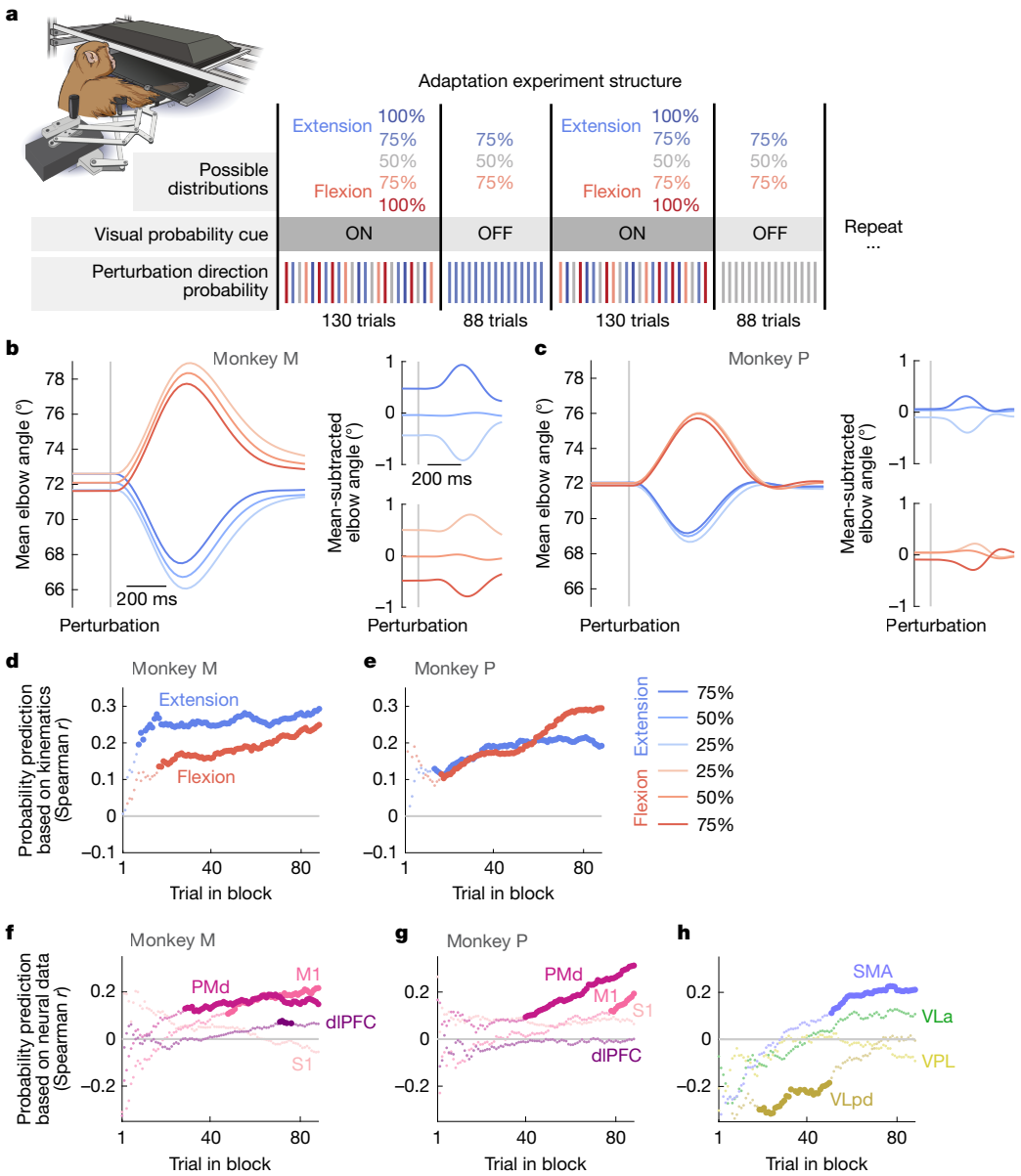


Fig. 4 | Only some areas represent sensory expectations accumulated from experience. a, Monkeys performed an adaptation variant of the main experiment in which trials alternated in block between randomly visually cued probabilities (as in the main experiment) and adaptation blocks without probability cue. Here, the perturbation probabilities were 75% extension, 50/50% or 75% flexion, fixed within each block. Image credit: Erika Woodrum, Erika Woodrum Art, Ltd. **b**,**c**, Mean elbow angle of each condition for trials with monkey M (**b**) and monkey P (**c**) within the adaptation blocks only. **d**,**e**, We aggregated all of the visually cued trials across sessions (for monkey M (**d**) and monkey P (**e**) and perturbation direction separately) and fit linear regressions to predict probability on the basis of the post-perturbation kinematics (shoulder and elbow velocity in the window 1–400 ms post-perturbation; lasso regularization parameter for least-squares linear regression selected using Bayesian optimization with 5-fold cross-validation). We used this fitted regression to predict the probability condition on adaptation trials and assessed performance as the Spearman correlation for trials at all points within an adaptation block. Correlations used all trials within a sliding window that included the 50 trials up to and including the current trial in block. Chance level calculations for remaining panels were made by randomly shuffling (10,000 iterations) the probability values of each trial for each point within a block and calculating the Spearman correlation with the true value. Correlations were considered significant (thick solid colours) if they exceeded the 99.9th percentile of the random distribution. **f**,**g**, Same analysis as in **d** (**f**) and **e** (**g**), but using neural population activity in the 300 ms prior to the perturbation to predict probability condition (linear regression with L2 penalty). **h**, Same analysis as in **g**, for additional areas in monkey P.

significant representation of probability condition 10–20 trials into the adaptation block. An optimal Bayesian integrator would reach 95% confidence in the most likely probability distribution after around 15 trials in the case in which each of the 3 distributions are equally likely, as was the case in our experiment.

To test which brain areas show a similar representation of sensory expectations in the visually cued and adaptation conditions, we repeated the previous regression analysis instead using populations of neurons in each area to predict probability, training the regression on the visually cued blocks and testing on the adaptation blocks (Fig. 4f,g). Across monkeys, we found that probability representations in the adaptation blocks were only found consistently in PMd and M1, starting earlier in PMd, suggesting that these areas either are involved in the accumulation of evidence forming particular sensory expectations, or receive this information from other unrecorded areas. In monkey P, despite some subcortical (VLa and VLpd) and cortical (SMA) areas showing probability representations during our visually cued experiment, only SMA and VLpd showed a clear representation of probability

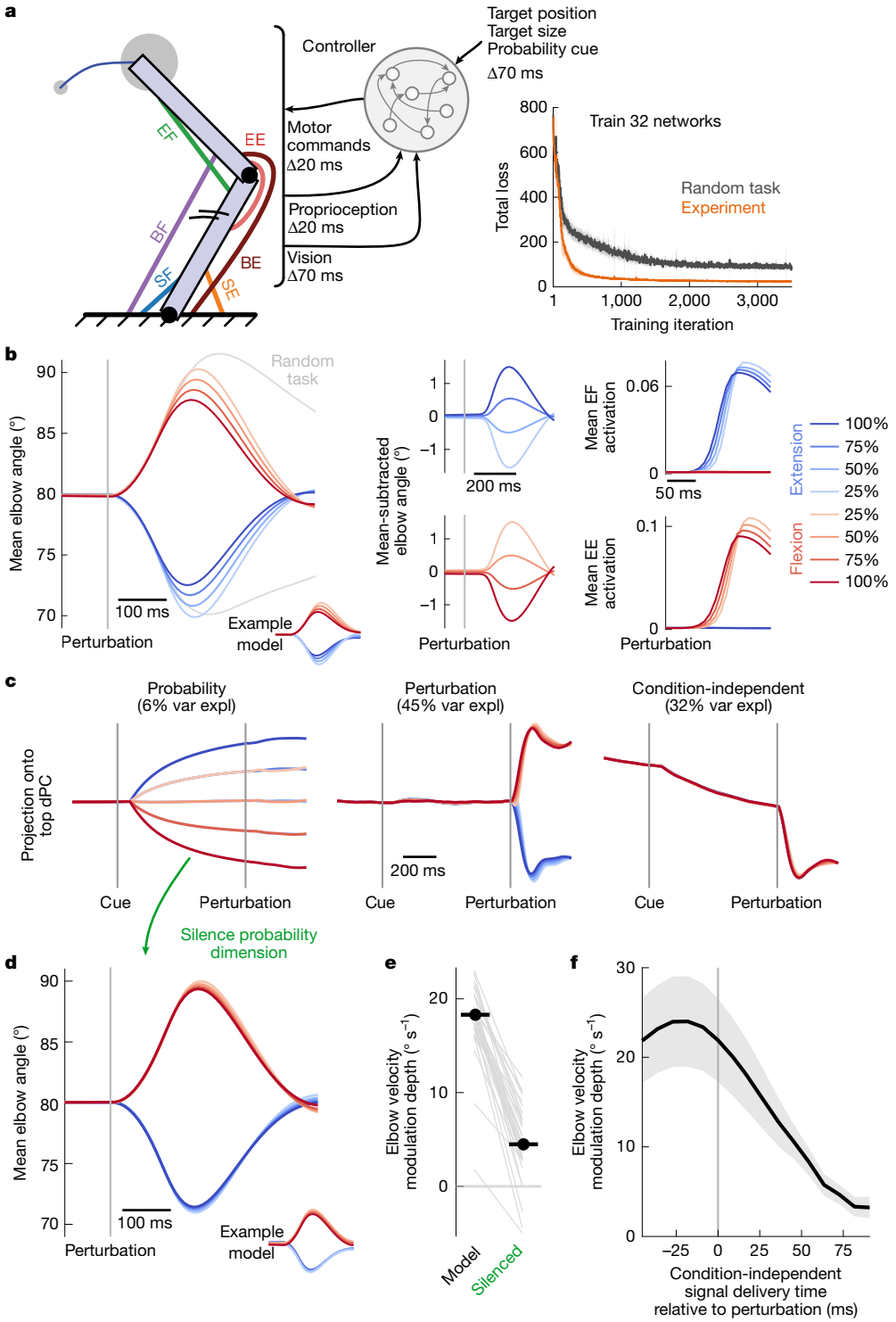


Fig. 5 | Models of closed-loop motor control develop sensory expectations dependent on feedback timing. a, We used an open source toolbox, MotorNet[3], to train 32 recurrent neural networks to control a biomechanical model of the arm during reaching, including realistic muscles, feedback and delays. A condition-independent pulse signalling the perturbation was provided along with proprioceptive feedback. Models were trained both on a random reaching task with targets throughout the workspace that had to be reached after a random mechanical perturbation, and a version of the experiment performed by humans and monkeys. BE, biceps extension; BF, biceps flexion; EE, elbow extension; EF elbow flexion; SE, shoulder extension; SF, shoulder flexion. Image adapted with permission from ref. 3, eLife. **b**, Left, average elbow kinematics of all models and an example model. Middle, the same kinematics with the mean activity subtracted. Right, average muscle activation of two simulated muscles, showing scaling of muscle activation within the long-latency reflex window (50–100 ms post-perturbation). **c**, We performed the same dPCA analysis as performed on neural data on the simulated neural activity (first dPC of an example model shown for each task factor). Var expl, variance explained. **d**, Mean elbow kinematics of models when the top probability dPC is silenced for the entire trial. **e**, Quantification of the reduction of kinematic modulation depth due to sensory expectations when silencing the top probability dPC. Each grey line corresponds to one model ($n = 32$ models). **f**, In all models, the time at which the condition-independent perturbation signal was provided to the network (original trained time was 20 ms after the perturbation) was manipulated in trained networks to examine its effect on the magnitude of sensory expectations in behaviour. Shaded error represents s.e.m. over models.

during the adaptation experiment. Of all areas examined, VLpd showed a significant representation of probability earliest in the block and did not show a significant representation at the end of the block. The sign of the significant correlation for VLpd was also reversed, suggesting that it may have a different role during visually cued trials and adaptation trials.

Together, these results suggest that some brain regions, most prominently PMd, show similar neural representations of probability during visually cued and adaptation trials, and that the earliest representations were present in the cerebello-thalamic inputs to cortex.

Sensory expectations emerge in models

Our results demonstrate the strong presence of sensory expectations in motor circuits. We next considered the conditions under which neural networks learn to represent sensory expectations and how such representations improve motor performance. To address this, we used our open source toolbox, MotorNet[3], to create closed-loop models of reaching in which recurrent neural networks actuate realistic muscles to control a two-link model of the arm and receive delayed sensory feedback (Fig. 5a and Methods).

Matching standard numbers of human participants included in behavioural experiments, we trained 32 networks on a random reaching task in which they produced reaches between random points in the workspace after receiving unpredictable perturbations delivered directly to the joints of the arm model. Paralleling the monkey's exposure to both everyday movements and the experimental task (Fig. 2), network training interleaved iterations of the random task and a version of the experimental task similar to the one the monkeys performed. During the experimental task, the networks also received a probability cue as in the human and monkey experiments. Delayed proprioceptive feedback allowed the model to detect displacement due to mechanical perturbations applied to the limb. Notably, mirroring what we observed empirically in early sensory areas (VPL and S1), we included a condition-independent perturbation signal as part of the proprioceptive feedback that signalled when a perturbation occurred but did not provide information about perturbation direction.

After training, model parameters were frozen, and they were tested on a version of the experimental task similar to Fig. 2. Of note, although the models were not trained to produce a specific movement trajectory or pattern of muscle activity, the models responded more quickly to perturbations that were more likely than to perturbations that were less likely (Fig. 5b), closely resembling humans and monkeys. The models learned this association between probability cues and perturbation probability by experiencing many perturbations paired with each visual cue, and the same pattern of results was obtained in networks with a one-hot input for visual cues (Extended Data Fig. 5), as well as networks reaching peripheral targets as in Fig. 1. Behavioural effects were due to scaling of muscle responses soon after perturbation onset, starting in an epoch akin to the long-latency reflex window due to the sensory delays introduced into our feedback loop (Fig. 5b, right).

Decomposing neural activity of the model in the same way that we did for monkey data using dPCA revealed very similar dimensions for probability, perturbation and condition-independent signals as observed in the empirical data (Fig. 5c). Notably, the geometry of sensory expectations was simple and similar to the empirical data, directly representing relative probability. To test whether this probability dimension was causally responsible for the behavioural effects that we observed, we eliminated all neural activity projecting into this dimension by subtracting the appropriate amount of neural activity from each neuron at each time point in the trial, without retraining the networks. Figure 5d demonstrates that when this dimension was eliminated the scaling of kinematic responses by probability was almost completely abolished. This effect was quantified by calculating the modulation depth due to probability during movement (Fig. 5e; maximum divergence between

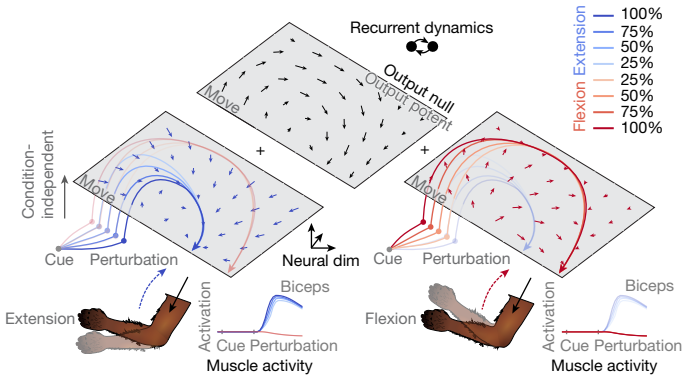


Fig. 6 | Schematic of the neural dynamics of sensory expectations. Motor cortical areas represent sensory expectations during preparation as the relative probability of events (in this study perturbation direction). Once an output-potent response is triggered by a condition-independent input, muscle activity initially reflects expectations about sensory inputs as the neural state evolves through a flow field determined by recurrent dynamics. As additional sensory information about the disturbance continues to arrive (in our case directional information about the mechanical perturbation), the influence of sensory feedback on the flow field pushes neural trajectories onto the path necessary to produce the muscle activity needed to counteract the actual disturbance. Dim, dimension. Image adapted with permission from ref. 3, eLife.

elbow kinematics for 100% flexion and 100% extension conditions in −100 to 300 ms around perturbation onset), which confirmed that the effect of sensory expectations on motor performance were almost completely driven by this dimension.

Finally, it was important to understand what constraints allowed the model to take advantage of probability information in its feedback responses. To investigate this question, we modified the timing of the condition-independent perturbation input by manipulating its latency from its original value of 20 ms post-perturbation, without retraining the networks (Fig. 5f). The effectiveness of sensory expectations decreased very quickly as the latency increased, showing essentially no effect once it was delayed more than about 50 ms, while reducing the latency increased the effectiveness of sensory expectations. Of note, when networks were trained from scratch without the condition-independent input to inform them of perturbation onset, they were eventually able to correct for perturbations, but they did not exhibit responses that scaled with expectations (Extended Data Fig. 6). That is, sensory expectations shape feedback responses when a perturbation is detected before sensory inputs resolve the ambiguity about what perturbation occurred and eliminate the need to respond on the basis of prior expectations.

Discussion

Our results demonstrate that humans and monkeys incorporate knowledge about future sensory inputs when preparing a movement and that this preparation improves their performance. Neural data show that information about sensory expectations is widespread across cortical and subcortical areas, generally following a simple neural geometry that directly represents the probability of each perturbation direction. A neural network trained to control a biomechanical model of the arm reveals that incorporating sensory expectations into movement preparation is advantageous when responding to such disturbances, provided that the perturbation is detected early enough to act on sensory expectations before incoming sensory information resolves the ambiguity about what perturbation occurred.

The neural representation of sensory expectations is consistent with our understanding of motor cortical control[19,43]. Motor cortex has an expansive ability to represent task variables in its preparatory

state[1], including prior information about goal location[44–47] and reward[48], but we demonstrate here that motor cortical areas directly represent expectations about sensory inputs during preparation (Fig. 6, output null trajectories). When disturbances do occur, responses are triggered by a condition-independent signal (with a similar profile to the condition-independent signal that precedes voluntary movements[37]) that quickly produces a muscle response (Fig. 6, output-potent trajectories). This muscle response proportionally reflects current expectations about perturbation direction (Fig. 6, muscle activity), similar to how goal-directed movements following perturbations reflect a continually updating movement plan[49]. As in self-initiated movements, neural activity initially evolves in the neural space on the basis of the flow field determined by recurrent dynamics. As sensory information about the actual disturbance (in our study, the perturbation direction) arrives, the modified flow field directs the neural activity towards the trajectory appropriate for the muscle activity necessary to counteract the actual perturbation.

In this framework, the straightforward representation of sensory expectations in the neural preparatory state can co-exist with dimensions related to other task factors such as goal location, a prediction that can be tested directly in future experiments. Another open question is to what extent eye movements have a role in these neural representations. Although we did not track eye movements, the widespread presence of sensory expectation signals, as well as the alignment between neural data in the visually cued and adaptation experiments, suggest that eye movements to visual cues cannot fully explain the present results.

In addition to the directional proprioceptive information transmitted through cortical areas following perturbations[28,29], we found a prominent condition-independent signal upon which feedback responses are likely to rely[31]. The vast majority of information in VPL, the thalamic nucleus projecting most strongly to the primary somatosensory cortex, was condition-independent, suggesting that this signal could have originated in the periphery, spinal cord or brainstem. One possibility is that cutaneous receptors, when aggregated together at the level of second-order neurons (cuneate nucleus), transiently signal the presence of a perturbation[50]. Pacinian corpuscles are a likely candidate, owing to their sensitivity to high frequency vibration and the fact that their receptive fields cover such large areas, giving them low directional resolution[51]. Another possibility is that fusimotor drive through gamma motor neurons increases muscle spindle sensitivity such that transient vibrations during perturbations in any direction produce a condition-independent signal through Ia afferents[52]. Determining the general function of this condition-independent signal and the circuit that constructs it is an important area for future work.

Our results have implications for predictive coding, which has been proposed as a widespread mechanism for increasing sensitivity to input across the brain[53–55]. We show that the early sensory areas investigated (VPL and S1) show very little predictive coding in our task, but very strong condition-independent signals. Predictive feedback in early sensory areas has been proposed as a mechanism to improve state estimation during movement in the presence of delays[56]. However, in our experiments, there were no ongoing movements when perturbations arrived and therefore no corollary discharge of motor commands to sensory areas. The presence of the rapid condition-independent signal discussed above probably eliminates the need for improving state estimation immediately following unexpected perturbations, and we see no evidence that sensory predictions are used in this case to increase input sensitivity. Furthermore, in a supplementary analysis, we found no evidence for a neural representation of unsigned prediction error, a signal involved in predictive coding (Extended Data Fig. 7).

Only a subset of areas showed sensory expectations when these expectations were accumulated over trials, namely areas involved with motor preparation and execution (SMA, PMd and M1) and VLpd in the thalamus, which receives primarily cerebellar input and outputs to SMA/PMd/M1 (ref. 41). The presence of probability information earliest in the cerebellar thalamus is noteworthy, as the cerebellum has been identified as a crucial component of state estimation during action[57–59], and muscle responses related to the expected duration of a mechanical perturbation are eliminated in capuchin monkeys during cooling of the dentate nucleus[60]. Our results suggest that the cerebellum may be necessary for initially forming expectations on the basis of experience (as in our adaptation experiment), but may have a different role when probabilistic information is extracted from sensory inputs (as in our visually cued experiment), or once stable sensory expectations have been established.

In conclusion, our findings reveal that motor circuit dynamics are strongly shaped by sensory expectations and demonstrate a novel computational role for preparatory activity in motor circuits[1]: modulating responses to sensory feedback.

Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41586-025-09690-9.

1. Churchland, M. M. & Shenoy, K. V. Preparatory activity and the expansive null-space. *Nat. Rev. Neurosci.* 25, 213–236 (2024).
2. Shenoy, K. V., Sahani, M. & Churchland, M. M. Cortical control of arm movements: a dynamical systems perspective. *Annu. Rev. Neurosci.* **36**, 337–359 (2013).
3. Codol, O., Michaels, J. A., Kashefi, M., Pruszynski, J. A. & Gribble, P. L. MotorNet, a Python toolbox for controlling differentiable biomechanical effectors with artificial neural networks. *eLife* **12**, RP88591 (2024).
4. Tanji, J. & Evarts, E. V. Anticipatory activity of motor cortex neurons in relation to direction of an intended movement. *J. Neurophysiol.* **39**, 1062–1068 (1976).
5. Churchland, M. M., Santhanam, G. & Shenoy, K. V. Preparatory activity in premotor and motor cortex reflects the speed of the upcoming reach. *J. Neurophysiol.* **96**, 3130–3146 (2006).
6. Messier, J. & Kalaska, J. F. Covariation of primate dorsal premotor cell activity with direction and amplitude during a memorized-delay reaching task. *J. Neurophysiol.* **84**, 152–165 (2000).
7. Weinrich, M., Wise, S. P. & Mauritz, K. H. A neurophysiological study of the premotor cortex in the rhesus monkey. *Brain* **107**, 385–414 (1984).
8. Kaufman, M. T., Churchland, M. M., Ryu, S. I. & Shenoy, K. V. Cortical activity in the null space: permitting preparation without movement. *Nat. Neurosci.* **17**, 440–448 (2014).
9. Li, N., Chen, T.-W., Guo, Z. V., Gerfen, C. R. & Svoboda, K. A motor cortex circuit for motor planning and movement. *Nature* **519**, 51–56 (2015).
10. Riehle, A. & Requin, J. Monkey primary motor and premotor cortex: single-cell activity related to prior information about direction and extent of an intended movement. *J. Neurophysiol.* **61**, 534–549 (1989).
11. Cisek, P. & Kalaska, J. F. Neural correlates of reaching decisions in dorsal premotor cortex: specification of multiple direction choices and final selection of action. *Neuron* **45**, 801–814 (2005).
12. Churchland, M. M., Afshar, A. & Shenoy, K. V. A central source of movement variability. *Neuron* **52**, 1085–1096 (2006).
13. Riehle, A. & Requin, J. The predictive value for performance speed of preparatory changes in neuronal activity of the monkey motor and premotor cortex. *Behav. Brain Res.* **53**, 35–49 (1993).
14. Churchland, M. M., Yu, B. M., Ryu, S. I., Santhanam, G. & Shenoy, K. V. Neural variability in premotor cortex provides a signature of motor preparation. *J. Neurosci.* **26**, 3697–3712 (2006).
15. Afshar, A. et al. Single-trial neural correlates of arm movement preparation. *Neuron* **71**, 555–564 (2011).
16. Michaels, J. A., Dann, B., Intveld, R. W. & Scherberger, H. Predicting reaction time from the neural state space of the premotor and parietal grasping network. *J. Neurosci.* **35**, 11415–11432 (2015).
17. Churchland, M. M. & Shenoy, K. V. Delay of movement caused by disruption of cortical preparatory activity. *J. Neurophysiol.* **97**, 348–359 (2007).
18. Li, N., Daie, K., Svoboda, K. & Druckmann, S. Robust neuronal dynamics in premotor cortex during motor planning. *Nature* **532**, 459–464 (2016).
19. Churchland, M. M. et al. Neural population dynamics during reaching. *Nature* **487**, 51–56 (2012).
20. Logiaco, L., Abbott, L. F. & Escola, S. Thalamic control of cortical dynamics in a model of flexible motor sequencing. *Cell Rep.* **35**, 109090 (2021).
21. Michaels, J. A., Schaffelhofer, S., Agudelo-Toro, A. & Scherberger, H. A goal-driven modular neural network predicts parietofrontal neural dynamics during grasping. *Proc. Natl Acad. Sci. USA* **117**, 32124–32135 (2020).
22. Todorov, E. & Jordan, M. I. Optimal feedback control as a theory of motor coordination. *Nat. Neurosci.* **5**, 1226–1235 (2002).

23. Crevecoeur, F. & Scott, S. H. Priors engaged in long-latency responses to mechanical perturbations suggest a rapid update in state estimation. *PLoS Comput. Biol.* **9**, e1003177 (2013).
24. Pruszynski, J. A. & Scott, S. H. Optimal feedback control and the long-latency stretch response. *Exp. Brain Res.* **218**, 341–359 (2012).
25. Hatsopoulos, N. G. & Suminski, A. J. Sensing with the motor cortex. *Neuron* **72**, 477–487 (2011).
26. Pruszynski, J. A. et al. Primary motor cortex underlies multi-joint integration for fast feedback control. *Nature* **478**, 387–390 (2011).
27. Evarts, E. V. & Tanji, J. Gating of motor cortex reflexes by prior instruction. *Brain Res.* **71**, 479–494 (1974).
28. Pruszynski, J. A., Omrani, M. & Scott, S. H. Goal-dependent modulation of fast feedback responses in primary motor cortex. *J. Neurosci.* **34**, 4608–4617 (2014).
29. Omrani, M., Murnaghan, C. D., Pruszynski, J. A. & Scott, S. H. Distributed task-specific processing of somatosensory feedback for voluntary motor control. *eLife* **5**, e13141 (2016).
30. Picard, N. & Smith, A. M. Primary motor cortical responses to perturbations of prehension in the monkey. *J. Neurophysiol.* **68**, 1882–1894 (1992).
31. Evarts, E. V. & Fromm, C. Sensory responses in motor cortex neurons during precise motor control. *Neurosci. Lett.* **5**, 267–272 (1977).
32. Wolpaw, J. R. Amplitude of responses to perturbation in primate sensorimotor cortex as a function of task. *J. Neurophysiol.* **44**, 1139–1147 (1980).
33. Reschechtko, S. & Pruszynski, J. A. Stretch reflexes. *Curr. Biol.* **30**, R1025–R1030 (2020).
34. Cheney, P. D. & Fetz, E. E. Corticomotoneuronal cells contribute to long-latency stretch reflexes in the rhesus monkey. *J. Physiol.* **349**, 249–272 (1984).
35. Pruszynski, J. A., Kurtzer, I., Lillicrap, T. P. & Scott, S. H. Temporal evolution of automatic gain-scaling. *J. Neurophysiol.* **102**, 992–1003 (2009).
36. Lara, A. H., Elsayed, G. F., Zimnik, A. J., Cunningham, J. P. & Churchland, M. M. Conservation of preparatory neural events in monkey motor cortex regardless of how movement is initiated. *eLife* **7**, e31826 (2018).
37. Kaufman, M. T. et al. The largest response component in the motor cortex reflects movement timing but not movement type. *eNeuro* **3**, ENEURO.0085-16.2016 (2016).
38. Trautmann, E. M. et al. Large-scale high-density brain-wide neural recording in nonhuman primates. *Nat. Neurosci.* **28**, 1562–1575 (2025).
39. Darian-Smith, C., Tan, A. & Edwards, S. Comparing thalamocortical and corticothalamic microstructure and spatial reciprocity in the macaque ventral posterolateral nucleus (VPLc) and medial pulvinar. *J. Comp. Neurol.* **410**, 211–234 (1999).
40. Horne, M. K. & Tracey, D. J. The afferents and projections of the ventroposterolateral thalamus in the monkey. *Exp. Brain Res.* **36**, 129–141 (1979).
41. Morel, A., Liu, J., Wannier, T., Jeanmonod, D. & Rouiller, E. M. Divergence and convergence of thalamocortical projections to premotor and supplementary motor cortex: a multiple tracing study in the macaque monkey: Thalamocortical connections of premotor cortex. *Eur. J. Neurosci.* **21**, 1007–1029 (2005).
42. Rouiller, E. M., Liang, F., Babalian, A., Moret, V. & Wiesendanger, M. Cerebellothalamocortical and pallidothalamocortical projections to the primary and supplementary motor cortical areas: a multiple tracing study in macaque monkeys. *J. Comp. Neurol.* **345**, 185–213 (1994).
43. Vyas, S., Golub, M. D., Sussillo, D. & Shenoy, K. V. Computation through neural population dynamics. *Annu. Rev. Neurosci.* **43**, 249–275 (2020).
44. Mauritz, K. H. & Wise, S. P. Premotor cortex of the rhesus monkey: neuronal activity in anticipation of predictable environmental events. *Exp. Brain Res.* **61**, 229–244 (1986).
45. Glaser, J. I., Perich, M. G., Ramkumar, P., Miller, L. E. & Kording, K. P. Population coding of conditional probability distributions in dorsal premotor cortex. *Nat. Commun.* **9**, 1788 (2018).
46. Rickert, J., Riehle, A., Aertsen, A., Rotter, S. & Nawrot, M. P. Dynamic encoding of movement direction in motor cortical neurons. *J. Neurosci.* **29**, 13870–13882 (2009).
47. Bastian, A., Schöner, G. & Riehle, A. Preshaping and continuous evolution of motor cortical representations during movement preparation. *Eur. J. Neurosci.* **18**, 2047–2058 (2003).
48. Smoulder, A. L. et al. A neural basis of choking under pressure. *Neuron* **112**, 3424–3433.e8 (2024).
49. Selen, L. P. J., Shadlen, M. N. & Wolpert, D. M. Deliberation in the motor system: reflex gains track evolving evidence leading to a decision. *J. Neurosci.* **32**, 2276–2286 (2012).
50. Johansson, R. S. & Flanagan, J. R. Coding and use of tactile signals from the fingertips in object manipulation tasks. *Nat. Rev. Neurosci.* **10**, 345–359 (2009).
51. Turecek, J. & Ginty, D. D. Coding of self and environment by Pacinian neurons in freely moving animals. *Neuron* **112**, 3267–3277.e6 (2024).
52. Dimitriou, M. & Edin, B. B. Human muscle spindles act as forward sensory models. *Curr. Biol.* **20**, 1763–1767 (2010).
53. Jiang, L. P. & Rao, R. P. N. Predictive coding theories of cortical function. in *Oxford Research Encyclopedia of Neuroscience* https://doi.org/10.1093/acrefore/9780190264086.013.328 (2022).
54. Richter, D., Kietzmann, T. C. & de Lange, F. P. High-level visual prediction errors in early visual cortex. *PLoS Biol.* **22**, e3002829 (2024).
55. Rouhani, N. & Niv, Y. Signed and unsigned reward prediction errors dynamically enhance learning and memory. *eLife* **10**, e61077 (2021).
56. Li, J. S., Sarma, A. A., Sejnowski, T. J. & Doyle, J. C. Internal feedback in the cortical perception-action loop enables fast and accurate behavior. *Proc. Natl Acad. Sci. USA* **120**, e2300445120 (2023).
57. Wolpert, D. M., Miall, R. C. & Kawato, M. Internal models in the cerebellum. *Trends Cogn. Sci.* **2**, 338–347 (1998).
58. Miall, R. C., Christensen, L. O. D., Cain, O. & Stanley, J. Disruption of state estimation in the human lateral cerebellum. *PLoS Biol.* **5**, e316 (2007).
59. Diedrichsen, J., Criscimagna-Hemminger, S. E. & Shadmehr, R. Dissociating timing and coordination as functions of the cerebellum. *J. Neurosci.* **27**, 6291–6301 (2007).
60. Hore, J. & Vilis, T. Loss of set in muscle responses to limb perturbations during cerebellar dysfunction. *J. Neurophysiol.* **51**, 1137–1148 (1984).

Methods

Human experiment

Participants. Twenty healthy individuals (13 male and 7 female participants, 18–35 years of age, 2 left-handed) took part in this experiment. All participants reported normal or corrected-to-normal vision and no history of neuromuscular impairments. Prior to data collection, all participants provided informed written consent. Participants were paid for their time and were able to withdraw from the study at any time. The study was approved by the Office of Research Ethics at the University of Western Ontario.

Apparatus. Participants were seated with their right arm in a KINARM robot exoskeleton (Fig. 1a, BKIN Technologies[61]), allowing flexion and extension movement of the shoulder and elbow joints in the horizontal plane. The robot can independently apply specific flexion or torques at these joints. The two segments of the exoskeleton, consisting of the upper arm and forearm, have three adjustable cuff sizes to match the dimensions of the participant's arm. Foam pads were inserted into any remaining space between the cuffs to ensure tight coupling of the limb to the applied torques. After adjustment of the robot, calibration was performed to align a real-time, 0.5 cm diameter cursor on the right index fingertip of each participant. The hand-position feedback and visual targets of the experiment were displayed in the same horizontal plane as the arm movement. These virtual-reality images were projected in front of participants at eye-level via an LCD monitor onto a semi-silvered mirror. Before initiating the experiment, an opaque blinder was installed beneath the mirror to occlude direct vision of the physical right arm during all trials. Kinematic data were sampled at 1,000 Hz.

Experimental procedure. Throughout the duration of the experiment, a constant background load of 1 N m extension torque was applied at the elbow, pre-exciting the flexor muscles. The use of a background load extending the elbow increases the stability and magnitude of flexor muscle responses[35,62]. To initiate each trial, participants moved their hand to a target (0.5 cm diameter) representing the external angles of 80° and 60° for the elbow and shoulder joints, respectively. As instructed, participants tried to exert the minimum force necessary to hold their arm at the home target without co-contraction of antagonistic muscles. After 300 ms in the home target, the goal target (3.5 cm diameter) appeared for a random period between 400–600 ms (Fig. 1b). The design of this task, including the large peripheral target, mirrors many previous experiments designed to investigate the role of goal information on muscle responses to mechanical perturbations[63]. The goal target was presented at a location that could be reached with a 10° pure elbow flexion from the home target. Then the arrow(s) indicating the probability of elbow perturbation direction appeared for a random period between 800–1,100 ms before the perturbation was applied at the elbow. All 5 probability cues were equally likely, and randomly selected from a pool of 880 trials. The arrows were created with areas directly proportional to the per cent probability. The perturbation was then applied (step torque of ±1 N m), which either flexed their elbow, moving their hand into the target, or extended their elbow, moving the hand away from the target. At the moment the perturbation was applied, the probability cues disappeared and visual feedback about hand location disappeared for 50 ms. Participants were instructed to move to the grey target once they felt the perturbations, and to do so in less than 700 ms. If this was achieved, the target changed from grey to green. However, if participants took more than 700 ms, the target changed from grey to red. This feedback was used to prompt participants to move quicker for the next trial if they moved too slowly. If participants moved off the home location prior to the perturbation, the trial was aborted and repeated later in the experiment. No restrictions were implemented on the trajectory of their arm movements. Regardless of green or red feedback, after holding their arm at the goal target for 400 ms, the torque is returned to the level corresponding to the constant background load. Participants then immediately moved to the home button to start the next trial.

Participants completed 49 practice trials, which were not included in the analysis. As part of the 880 trials, participants randomly received 10 trials of each condition as a catch trial. During catch trials, the cue appeared but the perturbation was never applied. The target automatically turned green after 2 s of holding the home target. Catch trials were used to ensure participants were not moving before the onset of the perturbation. Rest breaks were provided throughout each experiment at approximately 15–20 min intervals or when requested by the participants.

Electromyographic recording. The skin above the muscles of interest was scrubbed using a piece of gauze soaked with rubbing alcohol. The EMG electrodes (Delsys Bagnoli-8 system with DE-2.1 sensors) were coated with conductive gel (Chattanooga REF4248). The electrodes were taped to the skin surface above the belly of three right arm muscles: the short head of the biceps brachii, an elbow flexor; the brachioradialis, an elbow flexor; and the medial head of the triceps brachii, an elbow extensor. The electrodes were aligned parallel to the muscle fibres. A reference electrode was secured on the left clavicle of each participant. EMG signals were amplified with a gain of 1,000 and digitally sampled at 1,000 Hz. The collected EMG data was then bandpass filtered at 10–500 Hz using a zero-phase, second-order Butterworth filter and full-wave rectified.

Muscle activity of elbow flexors were normalized by their mean activity from the last 200 ms prior to perturbation onset across all trials. Muscle activity of the elbow extensor, medial tricep, was normalized to mean EMG activity during three special trials at the start of the experiment. These three trials totalled 11 s with a constant 1 N m elbow flexion torque.

Non-human primate experiments

Subjects. Two male rhesus macaques (monkey M, *Macaca mulatta*, 10 kg, 15 years old; monkey P, *M. mulatta*, 16 kg, 11 years old) participated in the study, which is the standard sample size for electrophysiological experiments in the field. No randomization or blinding was undertaken. All procedures described below were approved by the Institutional Animal Care and Use Committee at Western University (protocol 2022-028).

Experimental procedure. The design of the main monkey experiment closely mirrored the human experiment. Throughout the experiment, a constant background load of 0.02 N m extension torque was applied at the elbow. On each trial, monkeys waited with their fingertip in a central target (located under the fingertip when the shoulder and elbow angles were 32° and 72°, respectively; target size: 1.2 cm diameter). After a variable delay (600–800 ms), one of the five possible probability cues appeared randomly. In the monkey experiment, the probability arrows were coloured to further differentiate them (dark blue, 100% extension; light blue, 75% extension; white, 50% extension; light orange, 25% extension; dark orange, 0% extension). If at any point before the perturbation, the hand went outside the home target, the trial was aborted. Trials were excluded from analysis if at any point during the delay period hand velocity exceeded 0.5 cm s^{-1}. For monkey P, these trials were aborted in real-time, whereas for monkey M they were excluded from analysis. After a variable delay of 800–1,200 ms, monkeys received one of two unpredictable elbow perturbations (±0.2 N m step torque) which served as a go cue to compensate for the perturbation and return to the central target. For monkey M, at the time of perturbation onset all visual feedback was frozen until the hand returned to the goal target. For monkey P, all visual feedback was frozen for 150 ms after the perturbation. After returning to the central target and holding the hand there for 700 ms, a liquid reward was given. In both cases the

probability cues remained on until the end of the trial. In 10% of trials, after 1,200 ms no perturbation was applied and a liquid reward was given. In perturbation trials, the amount of liquid given at the end of the trial scaled with the speed of the return movement. Trials in which the time between the perturbation and the reward exceeded 1.2 s were excluded from analysis.

Electrophysiological recording. We performed high-density extracellular recordings using multiple Neuropixels probe versions (Neuropixels 1.0: 1 cm length, Neuropixels 1.0 NHP: 1 cm length, and Neuropixels 1.0 NHP: 4.5 cm length). After training on basic tasks, both monkeys were implanted with custom 3D printed titanium implants (accurate to 0.2 mm) that were designed to precisely conform to their individual skulls as determined by a model obtained using micro-computed tomography. Titanium implants were fixed in the skull using a variable number of titanium screws and included a built-in recording chamber and head post. Neural recording targets were identified by registering the computed tomography to a pre-surgery MRI (3D Slicer), and identifying the 3D location of each brain area by warping segmentations from a composite macaque atlas to the individual MRI of each animal (NMT v2 (refs. 64,65), CHARM[64,66] and SARM[67] atlases, see ref. 68 for additional thalamic parcellations). The use of skull conforming titanium implants allowed us to precisely plan recording trajectories to target desired structures. The precision of our implantation technique was confirmed post-mortem in monkey M to be accurate to within <0.5 mm on the cortical surface. For the S1 recordings in both monkeys and for the VPL recordings in monkey P, targeting was tested by applying small displacements to the elbow joint in a passive context and showing selective responses. The localization of VPL is further supported by rapid condition-independent perturbation responses not present in other thalamic nuclei (Extended Data Fig. 4). Finally, in a separate experiment in monkey P targeting of the medial geniculate nucleus was confirmed using passive listening to auditory tones, further confirming the validity of our targeting methods.

After monkeys were trained in the experiment, craniotomies were performed over the planned recording areas. In monkey M, a large craniotomy was performed to expose the entire recording area, while in monkey P, small 2.7 mm burr holes were drilled over recording sites as needed. In monkey M, a custom holder was designed (Neuronitek) for use with 1.0 cm Neuropixels to allow insertion through the dura using 2–4 mm retractable guide tubes and actuated with Narishige microdrives. In monkey P, we created a new design (Neuronitek) for use with the 4.5 cm NHP Neuropixels to allow insertion through the dura using 9 mm retractable guide tubes and actuated using a manual microdrive. For each recording configuration, we 3D printed a custom holder (Formlabs 3B+, Grey resin V4) that aligned the Neuropixels along a specific, pre-defined trajectory targeting the areas of interest. Recordings in S1 primarily targeted Brodmann area 1 and area 3b, although recordings in monkey P sampled more densely from area 3b, and partially from 3a. Recordings in M1 targeted a mixture of gyral and sulcal M1, with the majority coming from gyral M1.

Neural data processing. Neural data were recorded from Neuropixels probes using SpikeGLX. Neural data were processed using a custom processing pipeline (https://github.com/JonathanAMichaels/PixelProcessingPipeline). For monkey M, action potential stream data were first drift corrected using spike localization and decentralized registration[69,70] implemented in spikeinterface[71], which was able to accurately track vertical probe drift and correct it. Due to the large craniotomy, some of these recordings had large drift (0–250 μm). Neural data were then processed with Kilosort 2.0 (ref. 72) to further stabilize recordings during spike sorting. For monkey P, drift was minimal due to small craniotomies (drift 0–15 μm), so we immediately processed the data using Kilosort 4.0 (ref. 72), including built-in drift correction. Single neurons were considered successfully recorded if they were flagged by Kilosort as single neurons using default parameters, and if they were stably recorded for the duration of the recording. To determine whether neurons were properly isolated over the course of the recording, we generated the average firing rates of each neuron for each condition of the main experiment (8 conditions) divided up into 5 equal blocks of trials, additionally averaging across all time in each trial (200 ms before cue onset to 300 ms after perturbation onset), which yielded a matrix of 8 × 5 values for each neuron. We then calculated the mean index of dispersion for each neuron (variance over time block divided by mean over time block, averaged across conditions) to estimate how stable each neuron was over the course of the recording. It is important to note that this metric does not test neurons for tuning to the task, only for reliable responses over the course of the recording. Neurons with an index of dispersion below 2 were included in further analysis. The majority of neurons had an index of dispersion <1, and shifting this threshold ±1 did not affect results.

In general, no set of selection criteria can definitively classify single neurons from multi-unit activity, so as an additional control we examined whether a stricter inclusion criteria of only units with a peak absolute template amplitude in the range typically associated with single neurons (150–300 μV) changed our main findings. The dPCA results remained qualitatively unchanged despite removing 53–55% of neurons. In general, it is unlikely that the types of analyses presented in this work would be very sensitive to the accidental inclusion of multi-unit activity, since at the population level linear dimensionality reduction techniques find similar results regardless of whether or not data are spike sorted at all[73].

Demixed principal components analysis. Principal component analysis (PCA) is commonly employed to reduce the dimensionality of high-dimensional datasets by finding a low dimensional representation that captures large amounts of variance using independent linear combinations of neurons. For PCA, given a matrix of data X, where each row contains the average firing rates of one neuron for all times and task conditions, PCA finds an encoder F and an equivalent decoder D, which minimizes the loss function $L = ||X - FDX||^2$ under the constraint that the principal axes are normalized and orthogonal, and therefore $D = F^T$. Unfortunately, data that are represented in this way often heavily mix the effect of different task parameters between latent dimensions. We would like to extract dimensions that dissociate our specific task conditions. To achieve this, we performed dPCA with 20 latent dimensions[29] using freely available code: http://github.com/machenslab/dPCA. In contrast to PCA, dPCA seeks to explain marginalized variance with respect to our specific task variables (probability, perturbation, and time), instead of merely explaining total variance. Unlike PCA, dPCA utilizes a separate encoder and decoder, such that the loss being optimized was $L = \sum_\Phi L_\Phi = \sum_\Phi (||X_\Phi - F_\Phi D_\Phi X||^2 + \lambda_\Phi ||F_\Phi D_\Phi||^2)$, where X_Φ is the marginalization of the full data with respect to each of our task parameters of interest and the λ term is a regularization parameter, preventing overfitting. Marginalizations of X can be obtained by averaging over all parameters which are not being investigated and subtracting all simpler marginalizations. In our case the marginalizations of interest were probability × time, perturbation × time, time and probability × perturbation × time. The specific value of λ was determined using 5-fold cross-validation for each brain area in each monkey, allowing each factor to have a different value of λ_Φ.

dPCA requires data for all combinations of levels of each factor, which was not the case for our data, since in the 100% probability conditions the opposite perturbation (0% likely) never occurred. To handle this small amount of missing data, we used a technique proposed in the original dPCA paper and fit a generalized linear model to each neuron at each time point, using the task factors (probability and perturbation) as a design matrix. Using this fit, we generated surrogate data for the missing conditions to obtain a balanced design. In order to match trial-to-trial variability, firing rates included random Gaussian noise

that scaled with the standard error of each model coefficient. While this simulated data was used for fitting dPCA, in no case was it used during analysis or calculation of variance explained.

Motor control model. We trained a number of neural network models to control a biomechanical model of the arm by actuating simulated muscles during reaching using our previously developed open source toolbox, MotorNet[3]. For all models, the timestep size was 0.01 s, and we included a proprioceptive delay (20 ms), a visual delay (70 ms), and a muscle output delay (20 ms). We additionally included Gaussian noise at all time steps in the proprioceptive signal (s.d.: 10^{-3}), vision signal (s.d.: 10^{-3}), and in the muscle activation signal (s.d.: 10^{-4}). Effectors were actuated using numerical integration with the Euler method. The arm26 model used in this study is available under the RigidTendonArm26 Effector class. It is described briefly below. The skeleton of the arm26 models follow the formalization proposed in Mussa-Ivaldi et al.[74]. The full formalization of the Hill-type muscles can be found in Thelen[75], equations 1–7, and with the parameter values used in that study. When different parameters were provided for young and old individuals, the values for young individuals were used. In the RigidTendonArm26 class the moment arms are approximated as described in Kistemaker et al.[76], equations A10–A12.

Recurrent neural network architecture. All networks consisted of one layer of gated recurrent units (GRUs) with 256 units and standard activations (update/reset: sigmoid, candidate: tanh). Kernel and recurrent weights were initialized using Glorot initialization[77] and orthogonal initialization[78], respectively. At all time points we included gaussian noise in the candidate activation (before nonlinearity, s.d.: 10^{-3}). Biases were initialized at 0. Fifty per cent of GRUs (equivalent results if 100%) were connected to the output layer of one node per muscle with a sigmoid activation function. The output layer's kernel weights were initialized using Glorot normalization, and its bias was initialized at a constant value of −3. Because the output activation function is a sigmoid, this initial bias forces the output of the policy to be close to 0 at the start of initialization, ensuring a stable initialization state. Fifty per cent of GRUs (equivalent results if 100%) received task-related and feedback inputs and these units were non-overlapping with units connected to the output layer. As task-related inputs, networks received a delayed vector (70 ms delay) of (x, y) Cartesian coordinates for the start position and target position, target size, directional elbow perturbation probability (−1 to 1), and a binary cues indicating when the elbow probability cues was on, resulting in a 7-element input vector. Networks also received delayed feedback (20 ms delay) from the environment consisting of proprioceptive signals containing muscle length and velocity for each muscle, vision of the (x,y) position of the endpoint (70 ms delay), and a non-directional perturbation pulse (equal to one 20 ms after the perturbation, otherwise 0), resulting in a 15-element feedback vector.

Network training. Networks received interleaved training on a random reaching task and a probabilistic perturbation task. In the random reaching task, trials consisted of delayed reaches between random locations in the reachable workspace, where movement started after an unpredictable mechanical perturbation (random uniform −2 to 2 N m perturbations shoulder and elbow) and no probability cues were given. Target size was randomized (0–10 cm diameter). In the probabilistic perturbation task details closely matched the human and monkey experiments. The start/end location was 60 degrees shoulder and 80 degrees elbow angle, targets were 1.2 cm diameter, and perturbations were −1 or 1 N m elbow perturbations. In all training there was no background load and the randomized timing of cues was similar to the monkey experiment. Fifty per cent of trials were catch trials (no perturbation) to prevent unwanted premature movements. Each training iteration consisted of a batch of 64 trials, each 3 s long, and we used the Adam[79] optimizer with a learning rate of 3×10^{-3}. Each task was trained for 2,000 iterations.

Networks were optimized using a total loss that was a weighted sum of individual loss components, each addressing different aspects of the model's performance:

$$\begin{aligned}\mathcal{L} = {} & \lambda_{\text{Cartesian}} \frac{1}{BT} \sum\sum\sum |xy - \text{target}| \\ & + \lambda_{\text{muscle}} \frac{1}{BT} \sum\sum\sum F + \lambda_{\text{velocity}} \frac{1}{BT} \sum\sum\sum v^2 \\ & + \lambda_{\text{activity}} \frac{1}{BT} \sum\sum\sum h^2 + \lambda_{\text{spectral}} \frac{1}{BT} \sum\sum\sum (\Delta^2 h)^2 \\ & + \lambda_{\text{jerk}} \frac{1}{BT} \sum\sum\sum (\Delta^2 v)^2\end{aligned}$$

where B is the batch size, T is the total number of time steps in an episode, xy and target are the current and target cartesian endpoints, F is the force applied by all muscles, v is the Cartesian endpoint velocity, and h is the hidden activity of the network. Each component had a specific weight during training, specifically, $\lambda_{\text{Cartesian}} = 10^3$, $\lambda_{\text{muscle}} = 10^0$, $\lambda_{\text{velocity}} = 2 \times 10^2$, $\lambda_{\text{activity}} = 10^{-1}$, $\lambda_{\text{spectral}} = 10^4$, $\lambda_{\text{jerk}} = 10^6$. The specific values of each of these components was not critical to successful training, with a few caveats. The muscle force penalty was necessary to prevent the network from simply using high levels of force at all times, the spectral penalty was necessary to prevent networks from learning chaotic dynamics as a result of the delayed sensory feedback, and the jerk penalty sped up training by encouraging networks to respond robustly to mechanical perturbations. All networks were trained under the same loss function, but were initialized and trained under different random seeds, leading to differing initial network weights and simulated trials.

Assessing relative model contributions to neural geometry. Due to the multicollinearity of the four explanatory models we used in Extended Data Fig. 7, we used Shapley values derived from cooperative game theory[80,81] to estimate the true contribution of each model. To compute the Shapley value Φ_i for each predictor i, we use the formula:

$$\Phi_i = \sum_{S \subseteq N \setminus \{i\}} \frac{|S|!\,(|N| - |S| - 1)!}{|N|!} \cdot (f(S \cup \{i\}) - f(S)),$$

where N is the set of all predictors, $S \subseteq N \setminus \{i\}$ represents each possible subset of predictors that excludes i, $|S|$ is the size of subset S, and $f(S)$ is the performance metric (in our case, R^2) achieved by a model using only predictors in S. In all cases we used non-negative linear least-squares regression[82] (N-fold cross-validated across condition pairs) to fit the lower triangle the relational dissimilarity matrices (RDMs) of models to RDMs of neural data. RDMs were computed as the Euclidean distance between all pairs of conditions. For neural data, these RDMs were calculated in the high-dimensional neural space using the average firing rate of all neurons recorded within each brain area (pooled across sessions).

This formula evaluates the change in model performance, $f(S \cup \{i\}) - f(S)$, when predictor i is added to subset S, and weights each marginal contribution, ensuring equal representation of all subset sizes. Since Shapley values are calculated across a full set of subsets, each predictor's contribution was normalized to the model's total performance across all predictors. To match the amount of measurement noise between conditions, we took the most conservative approach and randomly downsampled all conditions to contain the same number of trials. We repeated the entire Shapley value calculation ten times with different random subsamples of trials and averaged the result.

For significance testing, after trial averaging we randomly shuffled the conditions for each neuron independently and repeated the entire Shapley value calculation 100 times, using this distribution as a null

distribution. A predictor's Shapley value was considered significant if it exceeded the 99th percentile of this null distribution.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

The neural data that support the findings of this study are openly available in the Dryad digital repository at https://doi.org/10.5061/dryad.0vt4b8hbr (ref. 83). Human data will be shared upon reasonable request by the corresponding author.

Code availability

Raw neural data were processed using a custom processing pipeline (https://github.com/JonathanAMichaels/PixelProcessingPipeline). Custom code for data analysis was written in MATLAB and Python and is available from the corresponding author upon request.

61. Scott, S. H. Apparatus for measuring and perturbing shoulder and elbow joint positions and torques during reaching. *J. Neurosci. Methods* **89**, 119–127 (1999).
62. Matthews, P. B. Observations on the automatic compensation of reflex gain on varying the pre-existing level of motor discharge in man. *J. Physiol.* **374**, 73–90 (1986).
63. Pruszynski, J. A., Kurtzer, I. & Scott, S. H. Rapid motor responses are appropriately tuned to the metrics of a visuospatial task. *J. Neurophysiol.* **100**, 224–238 (2008).
64. Jung, B. et al. A comprehensive macaque fMRI pipeline and hierarchical atlas. *Neuroimage* **235**, 117997 (2021).
65. Seidlitz, J. et al. A population MRI brain template and analysis tools for the macaque. *Neuroimage* **170**, 121–131 (2018).
66. Reveley, C. et al. Three-dimensional digital template atlas of the macaque brain. *Cereb. Cortex* **27**, 4463–4477 (2017).
67. Hartig, R. et al. The Subcortical Atlas of the Rhesus Macaque (SARM) for neuroimaging. *Neuroimage* **235**, 117996 (2021).
68. Hirai, T. & Jones, E. G. A new parcellation of the human thalamus on the basis of histochemical staining. *Brain Res. Rev.* **14**, 1–34 (1989).
69. Boussard, J., Varol, E., Lee, H. D., Dethe, N. & Paninski, L. Three-dimensional spike localization and improved motion correction for Neuropixels recordings. Preprint at *bioRxiv* https://doi.org/10.1101/2021.11.05.467503 (2021).
70. Varol, E. et al. in *ICASSP 2021–2021 IEEE International Conference on Acoustics, Speech and Signal Processing* 1085–1089 (IEEE, 2021).
71. Buccino, A. P. et al. SpikeInterface, a unified framework for spike sorting. *eLife* **9**, e61834 (2020).
72. Pachitariu, M., Sridhar, S., Pennington, J. & Stringer, C. Spike sorting with Kilosort4. *Nat. Methods* **21**, 914–921 (2024).
73. Trautmann, E. M. et al. Accurate estimation of neural population dynamics without spike sorting. *Neuron* **103**, 292–308.e4 (2019).
74. Mussa-Ivaldi, F. A., Hogan, N. & Bizzi, E. Neural, mechanical, and geometric factors subserving arm posture in humans. *J. Neurosci.* **5**, 2732–2743 (1985).
75. Thelen, D. G. Adjustment of muscle mechanics model parameters to simulate dynamic contractions in older adults. *J. Biomech. Eng.* **125**, 70–77 (2003).
76. Kistemaker, D. A., Wong, J. D. & Gribble, P. L. The central nervous system does not minimize energy cost in arm movements. *J. Neurophysiol.* **104**, 2985–2994 (2010).
77. Glorot, X. & Bengio, Y. Understanding the difficulty of training deep feedforward neural networks. In *Proc. 13th International Conference on Artificial Intelligence and Statistics* 249–256 (JMLR, 2010).
78. Hu, W., Xiao, L. & Pennington, J. Provable benefit of orthogonal initialization in optimizing deep linear networks. Preprint at https://doi.org/10.48550/arXiv.2001.05992 (2020).
79. Kingma, D. P. & Ba, J. Adam: A method for stochastic optimization. Preprint at https://doi.org/10.48550/arXiv.1412.6980 (2014).
80. Scott, M. & Su-In, L. A unified approach to interpreting model predictions. *Adv. Neural Inf. Process. Syst.* **30**, 4765–4774 (2017).
81. Shapley, L. S. in *Contribution to the Theory of Games* (eds Kuhn, H. & Tucker, A.) 307–317 (Princeton Univ. Press, 1953).
82. Diedrichsen, J. & Kriegeskorte, N. Representational models: a common framework for understanding encoding, pattern-component, and representational-similarity analysis. *PLoS Comput. Biol.* **13**, e1005508 (2017).
83. Michaels, J. A. & Pruszynski, J. A. Data from: Sensory expectations shape neural population dynamics in motor circuits [Dataset]. *Dryad* https://doi.org/10.5061/dryad.0vt4b8hbr (2025).

Acknowledgements Experiments were funded by CIHR Operating Grants (Foundation Grant to J.A.P.: 353197; Project Grant to J.A.P. and J.D.: PJT-175010), the Simons Foundation (via the Simons-Emory International Consortium on Motor Control), and the Azrieli Foundation (via the Collaboration on Motor Planning, Execution and Resilience). J.A.M. was supported by a Banting Postdoctoral Fellowship, a Vector Institute Postgraduate Affiliation, and a BrainsCAN Postdoctoral Fellowship (CFREF). J.A.P. was supported by the Canada Research Chair programme.

Author contributions J.A.M. and J.A.P. conceptualized the project. J.A.M., J.Z., J.W. and J.A.P. performed the human experiments. J.A.M., M.K., R.K., J.C.L. and J.A.P. performed the monkey experiments. J.A.M., M.K. and J.A.P. performed neural data analysis. J.A.M., O.C. and J.A.P. performed neural network simulations. J.A.M., P.L.G., J.D. and J.A.P. wrote the manuscript. All authors edited and approved the manuscript.

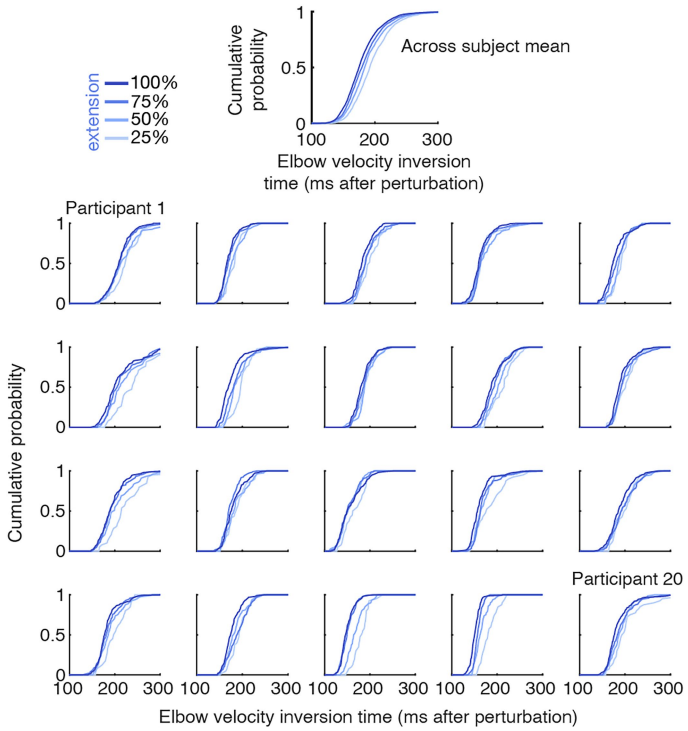
Competing interests The authors declare no competing interests.

Additional information

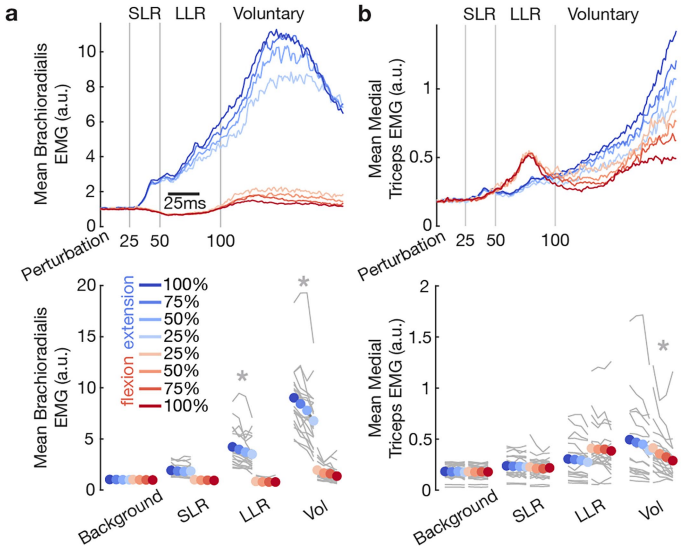
Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41586-025-09690-9.

Correspondence and requests for materials should be addressed to Jonathan A. Michaels or J. Andrew Pruszynski.

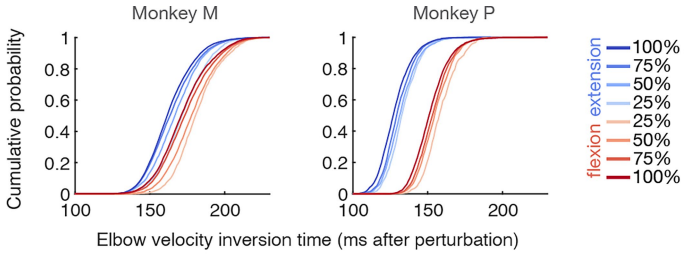
Peer review information *Nature* thanks Aaron Batista, who co-reviewed with Tapas Arakeri and Raeed Chowdhury, and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer review reports are available.



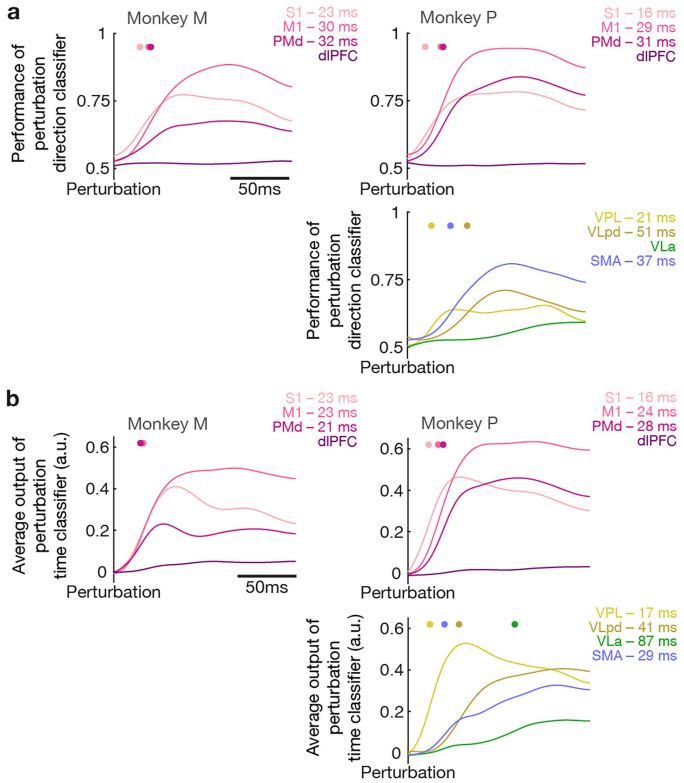
Extended Data Fig. 1 | Elbow velocity inversion time distributions in human participants. Visualization of the distribution of elbow velocity inversion times (i.e. time of maximum elbow joint excursion) relative to perturbation onset across all human participants, separated by probability condition. Only the extension perturbation conditions are shown, since flexion perturbations brought the hand into the target and did not require strong behavioral responses.



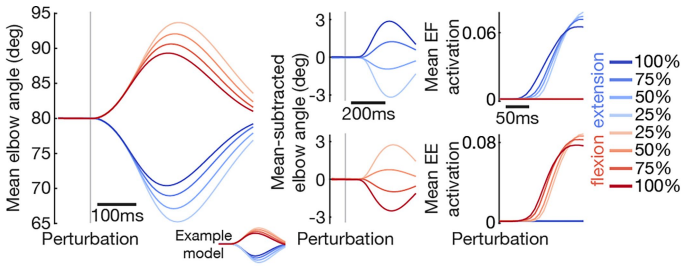
Extended Data Fig. 2 | Perturbation responses in human brachioradialis and medial triceps. a, In elbow extension conditions, no significant differences were found between probabilities in the background epoch (−200-0 ms before perturbation) of the brachioradialis ($F(3,19) = 1.57$, $p = 0.21$) or in the short latency (SLR, 20–50 ms) response ($F(3,19) = 1.73$, $p = 0.17$). In contrast, there was a significant effect of probability on EMG activity in the long latency epoch (LLR, 50–100 ms) of brachioradialis ($F(3,19) = 7.04$, $p = 0.0004$) and during the voluntary epoch (100–150 ms) of brachioradialis ($F(3,19) = 32.62$, $p < 0.0001$). **b**, In elbow flexion conditions, no significant differences were found between probabilities in the background epoch (−200-0 ms before perturbation) of the medial triceps ($F(3,19) = 0.31$, $p = 0.82$), in the short latency (SLR, 20–50 ms) response ($F(3,19) = 1.30$, $p = 0.28$), or in the long latency epoch (LLR, 50–100 ms) of the medial triceps ($F(3,19) = 0.99$, $p = 0.40$). In contrast, there was a significant effect of probability on EMG activity during the voluntary epoch (100–150 ms) of the medial triceps ($F(3,19) = 10.07$, $p < 0.0001$).



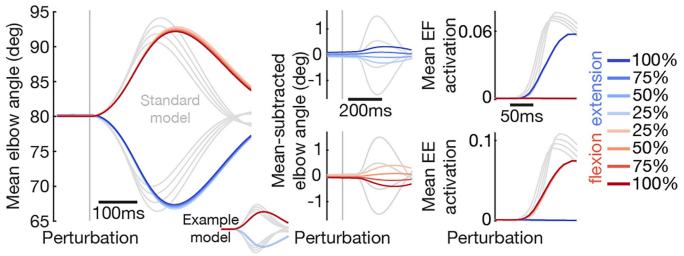
Extended Data Fig. 3 | Elbow velocity inversion times distributions in monkeys. Visualization of the distribution of elbow velocity inversion times (measured as ms after perturbation time) across all monkey sessions, separated by probability condition. All neural recording sessions within each monkey were pooled together.



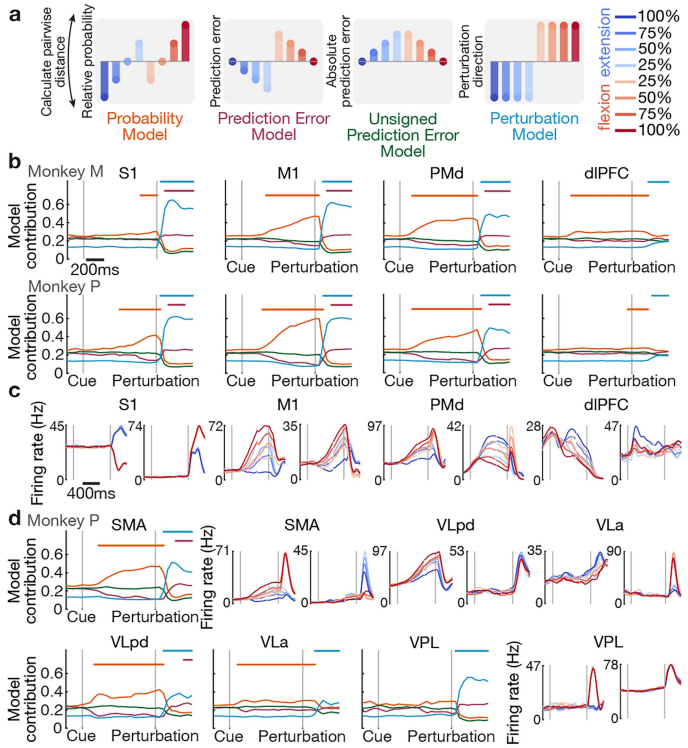
Extended Data Fig. 4 | Timing of perturbation direction detection and perturbation detection signals across areas. a, For each brain area, we analyzed neural population data to determine the time at which the perturbation direction could be detected. For each recording session, we fit a classifier to distinguish what perturbation was applied in all trials within a session (SVM, 5-fold cross-validated, 0 to 150 ms relative to perturbation onset). We analyzed the classifier performance across all trials within each area, smoothed the result (20 Hz low-pass 4th-order zero-phase butterworth), and determined the moment of maximum velocity in the classifier performance (colored dots above each plot and corresponding text). **b**, For each brain area, we analyzed neural population data to determine the time at which the perturbation could be detected independent of the perturbation direction. For each recording session, we fit a classifier to distinguish pre- and post-perturbation times across all trials within a session (SVM, 5-fold cross-validated, −150 to 150 ms relative to perturbation onset). We averaged the output of the classifier across all trials within each area, smoothed the result (20 Hz low-pass 4th-order zero-phase butterworth), and determined the moment of maximum velocity in the classifier output (colored dots above each plot and corresponding text). The classifier output is normalized to zero at the moment of perturbation for visualization purposes.



Extended Data Fig. 5 | Models with one-hot probability cue inputs develop sensory expectations. We trained 32 recurrent neural networks to control a biomechanical model of the arm during reaching, including realistic muscles, feedback, and delays, but using one-hot inputs to represent each probability cue (separate input channel for each cue) instead of the direct probability representation used in Fig. 6. Average kinematics and muscle responses were virtually identical to the main results of Fig. 6.



Extended Data Fig. 6 | Models without a condition-independent perturbation signal don't express sensory expectations. We trained 32 recurrent neural networks to control a biomechanical model of the arm during reaching, including realistic muscles, feedback, and delays, but omitting the condition-independent perturbation pulse included in the results of Fig. 6. Average kinematics of models performing the experiment did not scale with probability, nor did muscle activation within the long-latency reflex window (50–100 ms post-perturbation). The results of Fig. 6 are outlined in gray.



Extended Data Fig. 7 | Probability representations are widespread during preparation and are replaced by prediction error and perturbation direction representations during movement. a, We compared the ability of four potential models to explain neural population geometry by predicting the euclidean relational dissimilarity matrices (RDMs) between all pairs of conditions in the neural population space as a linear combination of model RDMs using non-negative linear least-squares regression. Neural RDMs were calculated in the high-dimensional neural space, while model RDMs were calculated in the 1-dimensional space of each model. **b**, Each model's unique contribution to explaining neural geometry was estimated using Shapley values. Fits were considered significant (solid bars above plots) if they exceeded the 99th percentile of the null distribution generated by randomly shuffling conditions for each neuron independently and repeating the Shapley value calculation 100 times. **c**, Example mean firing rates for single neurons across areas. **d**, Model fitting results and example neurons for medial wall and thalamic areas in Monkey P. In almost all areas investigated, the neural population data were organized such that the 100% probability conditions of opposite perturbation directions were the farthest away from each other in neural space and the other probability conditions were positioned in a graded fashion in between these extremes (a, Probability Model), indicating a direct representation of relative probability. However, predictive coding theories predict that some components of the neural response would represent how incongruous perturbations are with expectations (a, Prediction Error Model), scaling directionally based on how much the delivered perturbation deviated from expectation. Similarly, these theories predict the presence of a signal representing how surprising a given perturbation was (a, Unsigned Prediction Error Model), as both of these prediction error-related models have been shown to be relevant for learning and memory. Responses could also have differed purely based on actual perturbation direction (a, Perturbation Model). To test the relative contribution of these models to explaining neural data, we converted all models and neural population data into euclidean distance matrices between all pairs of conditions (relational dissimilarity matrices, RDMs) to represent the geometry of each model and the neural population. Disentangling the contribution of each of these models to observed neural activity is challenging due to multicollinearity that exists between models. To overcome this, we used a concept from cooperative game theory to estimate the contribution of each model. We used non-negative linear least-squares regression to predict the RDM of the neural population at each time point based on linear combinations of our model RDMs. Importantly, we exhaustively fit every possible combination of models, allowing us to calculate Shapley values, which estimate the true contribution of each model to explaining neural data (see Methods). During the preparatory period only the Probability Model was able to significantly fit the data, and this effect was widespread across cortical areas (b). The earliest and strongest representation was in PMd, while the latest and weakest was in S1. Probability representations collapsed dramatically within the 50–100 ms after the perturbation, and in many areas were replaced immediately by the Perturbation Model and the Prediction Error Model. Although perturbation direction representations were far more dominant, prediction errors were also reliably and significantly present, and can be seen in some example single neurons (c, e.g. 4th neuron from the left). Perhaps surprisingly, in no case did the Unsigned Prediction Error Model ever significantly explain neural geometry, indicating that responses directly related to surprise were not present. In Monkey P, we performed the same analysis for additional medial wall and thalamic areas (d). Significant linear representations of probability were also widespread during preparation in these areas, with the exception of VPL. Most of these areas also showed a significant representation of the Perturbation Model and the Prediction Error Model, while no area showed a significant representation of the Unsigned Prediction Error Model.

Extended Data Table 1 | Recording session information

Monkey M

Session	Recorded neurons	Successful trials	Experiment Type
06-Apr-22	M1: 26	783	Random
07-Apr-22	PMd: 62	845	Random
20-Apr-22	PMd: 32	1375	Random
22-Apr-22	M1: 134	1613	Random
28-Apr-22	S1: 50	1524	Random
29-Apr-22	M1: 82	1770	Random
04-May-22	PMd: 73	1421	Random
05-May-22	PMd: 32	1254	Random
06-May-22	S1: 39	1562	Random
12-May-22	PMd: 57	1657	Random
13-May-22	S1: 42	1691	Random
16-May-22	M1: 87	1932	Random
18-May-22	M1: 59, PMd: 133	1735	Random
25-May-22	PMd: 39	1855	Random
26-May-22	S1: 35, PMd: 219	1653	Random
27-May-22	S1: 20, PMd: 89	1699	Random
14-Dec-22	PMd: 102, dlPFC: 299	964	Random + Adaptation
15-Dec-22	PMd: 7, dlPFC: 347	1347	Random + Adaptation
02-Feb-23	PMd: 36, dlPFC: 128	1078	Random + Adaptation
03-Feb-23	PMd: 16, dlPFC: 309	1755	Random + Adaptation
09-Feb-23	PMd: 18, dlPFC: 24	910	Random + Adaptation
10-Feb-23	PMd: 37, dlPFC: 111	1344	Random + Adaptation
14-Feb-23	PMd: 2, dlPFC: 46	959	Random + Adaptation
15-Feb-23	PMd: 10, dlPFC: 89	1138	Random + Adaptation
16-Feb-23	M1: 60	1424	Random + Adaptation
17-Feb-23	S1: 81	1724	Random + Adaptation
24-Feb-23	S1: 136, M1: 132	1077	Random + Adaptation
01-Mar-23	S1: 116, M1: 207	1301	Random + Adaptation
03-Mar-23	PMd: 69	1719	Random + Adaptation

Monkey P

Session	Recorded neurons	Successful trials	Experiment Type
17-Jan-24	M1: 286	735	Random
18-Jan-24	M1: 283	928	Random
19-Jan-24	M1: 230	969	Random + Adaptation
23-Jan-24	PMd: 333	969	Random + Adaptation
24-Jan-24	PMd: 333	1126	Random + Adaptation
26-Jan-24	PMd: 175	1200	Random + Adaptation
29-Jan-24	PMd: 256	1419	Random + Adaptation
30-Jan-24	PFC: 231	1157	Random + Adaptation
31-Jan-24	dlPFC: 306	1347	Random + Adaptation
01-Feb-24	PMd: 336, dlPFC: 219	1221	Random + Adaptation
02-Feb-24	dlPFC: 329, VLA: 243	1231	Random + Adaptation
05-Feb-24	dlPFC: 342, VLA: 254	1262	Random + Adaptation
06-Feb-24	S1: 113	1539	Random + Adaptation
08-Feb-24	S1: 23, M1: 192	1342	Random + Adaptation
09-Feb-24	S1: 102, VLpd: 204	1413	Random + Adaptation
19-Feb-24	SMA: 157	688	Random + Adaptation
20-Feb-24	SMA: 95, preSMA: 92	1068	Random + Adaptation
21-Feb-24	SMA: 161, preSMA: 151	1615	Random + Adaptation
28-Feb-24	VPL: 131	1471	Random + Adaptation
29-Mar-24	M1: 164, PMd: 196	1120	Random + Adaptation

Date, number of recorded neurons from each area, number of successful trials, and experiment type for all individual recording sessions across both monkeys.

nature portfolio

Corresponding author(s): J. Andrew Pruszynski, Jonathan A. Michaels

Last updated by author(s): Jan 23, 2025

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☐	☒	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	Behavioral data collected using KINARM Dexterit-E (v3.5-3.9) Neural data collected using Spike-GLX (v3.0) NMT v2 Macaque Atlas was used for planning electrode trajectories in combination with pre-acquired MRI and CT. Further details regarding data collection are described in the Methods section of the manuscript.
Data analysis	MR and CT images were processed and co-registered using 3D Slicer (v5.2) Spike sorting was completed offline using either decentralized drift correction (within spikeinterface) and Kilosort 2.0, or using Kilosort 4.0. Neural network modeling used the MotorNet toolbox (v0.2.0). Custom Matlab (v2023b) and Python (v3.9) code was used for the analysis of all other data and is described in the Methods section of the manuscript.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

The neural data that support the findings of this study are openly available in the Dryad digital repository at https://doi.org/10.5061/dryad.0vt4b8hbr. Human data will be shared upon reasonable request by the corresponding author.

Research involving human participants, their data, or biological material

Policy information about studies with human participants or human data. See also policy information about sex, gender (identity/presentation), and sexual orientation and race, ethnicity and racism.

Reporting on sex and gender	Twenty healthy individuals (13 males and 7 females, aged 18-35 years, 2 left handed) participated in our human experiment. No analyses related to sex were undertaken, and no information about gender was collected.
Reporting on race, ethnicity, or other socially relevant groupings	No information regarding race, ethnicity, or other groups was collected.
Population characteristics	Participants were healthy individuals aged 18-35. All participants reported normal or corrected-to-normal vision and no history of neuromuscular impairments.
Recruitment	Participants were recruited through word of mouth and an internal university experimental subject pool. The experiment addresses low-level motor functions that are unlikely to be affected by a less diverse participant pool, although this question is a topic of future work.
Ethics oversight	Participants were paid for their time and were able to withdraw from the study at any time. The study was approved by the Office of Research Ethics at the University of Western Ontario.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	We analyzed data from two non-human primates, collected over dozens of sessions per animal. In non-human primate research, a few sessions across two animals is generally considered sufficient when results are robust across animals. We analyzed data from 20 human participants. No sample size calculation was performed, although this number of participants falls within the typical range for similar studies.
Data exclusions	No data were excluded from analysis.
Replication	All primary reported behavioral and neural results were independently replicated in each animal by pooling neural data across the many recording sessions of each animal. Human participant behavioral and physiological results were confirmed at the population level. No additional replications were undertaken.
Randomization	All subjects (either human or non-human primate) experienced the same experimental design and trials were randomized.
Blinding	All subjects (either human or non-human primate) experienced the same experimental design and trials were randomized. Experimenters did not have influence over presented conditions.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☐	☒ Animals and other organisms
☒	☐ Clinical data
☒	☐ Dual use research of concern
☒	☐ Plants

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☐	☒ MRI-based neuroimaging

Animals and other research organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research, and Sex and Gender in Research

Laboratory animals	Two male rhesus macaques (Monkey M, Macaca mulatta, 10 kg, 15 years old; Monkey P, Macaca mulatta, 16 kg, 11 years old) participated in the study.
Wild animals	No wild animals were used in this study.
Reporting on sex	Two males were used in the study. Sex was not considered in the study design.
Field-collected samples	No field-collected samples were used in this study.
Ethics oversight	All procedures described were approved by the Institutional Animal Care and Use Committee at Western University (Protocol #2022-028).

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Plants

Seed stocks	N/A
Novel plant genotypes	N/A
Authentication	N/A

Magnetic resonance imaging

Experimental design

Design type	N/A
Design specifications	N/A
Behavioral performance measures	N/A

Acquisition

Imaging type(s)	T1 Structural
Field strength	7T
Sequence & imaging parameters	Slice thickness = 0.5mm. Orientation = right-anterior-superior.
Area of acquisition	whole brain

Diffusion MRI ☐ Used ☒ Not used

Preprocessing

Preprocessing software	MR and CT images were processed and co-registered using Slicer3D (v5.2)
Normalization	The AFNI @animal_warper function was used for normalization. T1 sequences were volume normalized using a non-linear approach.
Normalization template	Macaque NMT v2
Noise and artifact removal	N/A
Volume censoring	N/A

Statistical modeling & inference

Model type and settings	N/A
Effect(s) tested	N/a

Specify type of analysis: ☒ Whole brain ☐ ROI-based ☐ Both

Statistic type for inference (See Eklund et al. 2016)	N/A
Correction	N/A

Models & analysis

n/a	Involved in the study
☒	☐ Functional and/or effective connectivity
☒	☐ Graph analysis
☒	☐ Multivariate modeling or predictive analysis