Cortical network and projection neuron types that articulate serial order in a skilled motor behavior

Yi Li [1,2], Xu An [1,2], Yongjun Qian [1,2], X. Hermione Xu [3], Shengli Zhao [1], Hemanth Mohan [1,2], Ludovica Bachschmid-Romano [1], Nicolas Brunel [1], Ian Q. Whishaw [4], Z. Josh Huang [1,2,3,5]

[1] Department of Neurobiology, Duke University, Durham, NC 27710, USA
[2] Cold Spring Harbor Laboratory, Cold Spring Harbor, NY 11724, USA
[3] Department of Biomedical Engineering, Duke University, Durham, NC 27710, USA
[4] Department of Neuroscience, Canadian Centre for Behavioural Research, University of Lethbridge, Lethbridge, AB, T1K 3M4, Canada
[5] Correspondence: josh.huang@duke.edu

ABSTRACT

Skilled motor behaviors require orderly coordination of multiple constituent movements with sensory cues towards achieving a goal, but the underlying brain circuit mechanisms remain unclear. Here we show that target-guided reach-grasp-to-drink (RGD) in mice involves the ordering and coordination of a set of forelimb and oral actions. Cortex-wide activity imaging of multiple glutamatergic projection neuron (PN) types uncovered a network, involving the secondary motor cortex (MOs), forelimb primary motor and somatosensory cortex, that tracked RGD movements. Photo-inhibition highlighted MOs in coordinating RGD movements. Within the MOs, population neural trajectories tracked RGD progression and single neuron activities integrated across constituent movements. Notably, MOs intratelencephalic, pyramidal tract, and corticothalamic PN activities correlated with action coordination, showed distinct neural dynamics trajectories, and differentially contributed to movement coordination. Our results delineate a cortical network and key areas, PN types, and neural dynamics therein that articulate the serial order and coordination of a skilled behavior.

Keywords: serial order, movement coordination, cortical network, motor cortex, neuron type, neural dynamics.

INTRODUCTION

To survive and thrive in their ethological niche, animals deploy skilled motor behaviors involving the orderly coordination of multiple movements cued by concurrent sensory streams to achieve an intended goal. For example, feeding behaviors in rodents and primates often consist of reaching for a food item with the arm, grasping food with the hand, and withdrawing hand to the mouth to eat [1]. While reach and grasp are allocentric movements that coordinate the arm and hand towards the food, withdraw is an egocentric movement that coordinates the hand with orofacial actions to eat [2]; thus a set of elemental actions are sequentially and continuously arranged with sensory feedback to compose a skillful natural behavior. The view of behavior as hierarchically organized and involving dynamic unfolding of serially ordered constituent movements was championed by Lashley [3] and has since been thoroughly corroborated [4], but the underlying brain circuit mechanisms, especially the role of cerebral cortex, remains poorly understood.

Beginning with the pioneering work of Evarts [5], motor neuroscience has focused on exploring the cortical encoding and control of relatively simple forelimb movements of the reach or grasp [5-9]. Studies have also examined the role of motor cortex in the sequential execution of individual movements of the arm [10-13] or tongue [14], i.e. “discrete action sequences” [15]. However, much less effort has been directed toward studying the ordering and coordination of “sequential continuous actions” across the body that compose an integrated and unified ethological behavior [15,16].

In exploring the cortical control of arm movements, most previous studies have focused on the broadly defined primary and higher order motor cortex [5,10,17-22] despite the recognition that sensory feedback are inherent and essential components of all skilled behaviors [23,24]. Thus a more holistic view of the extended cortical networks in the control and coordination of more complex motor behaviors has yet to be established. Within the motor areas, although decades of research has explored whether and how activities of individual or populations of neurons represent movement [5-9,25,26], as yet little consensus has been achieved on the basic response and encoding properties of cortical neurons [27,28]. As an alternative to this representational perspective, the dynamical systems approach seeks to understand how the dynamics of cortical neural populations produce the temporal patterns needed to drive movement [29], and significant experimental and computational work has identified rich structures within the coordinated activity of neural populations across motor areas and species [29-34]. Nevertheless, despite these seminal advances, a fundamental limitation in previous studies is the difficulty to resolve the identity of recorded neurons or neural populations at the level of molecular and projection defined neuron types, the basic elements of neural circuit [35]. As a consequence, it has been difficult to systematically identify potential neuronal subsets within the heterogeneous population that might encode movement parameters; it is also unclear how neural population dynamics are implemented at the cellular level, how they flow within neural circuits, and how they are transmitted to drive behavior, including the serial order and coordination of actions in behavior.

Laboratory mice display skilled motor behaviors such as reach, grasp, and manipulate to consume food and water [2,36,37]. Recent technical advances in machine learning-powered behavior tracking and analysis [38], wide-field imaging of neural activities [39-41], and cell-type targeted recording and manipulation enable a multi-faceted and integrated analysis of the neural circuit basis of these behaviors [42]. To date, most neural circuit studies in mice have focused on movement parts, e.g. the forelimb [40,43-48] or orofacial movement [49-55], but not their serial order and coordination across body as an integrated behavior (but see [36,56]). Here, through high-resolution quantitative analysis of a chemosensory-guided reach and grasp to drink behavior (RGD) [45], we have revealed the serial ordering and coordination of a set of forelimb (aim, advance, supinate, grasp, withdraw) and oral (mouth open, lick) actions to retrieve water to drink. Combining systematic cell-type resolution wide-field imaging, optogenetic manipulation, single neuron recording, and neural population dynamics analysis, we identified a dynamic cortical network that track RGD movement progression and uncovered the key role of MOs in the orderly coordination of forelimb and orofacial actions. We further discovered that activities of distinct PN types within MOs correlate with specific action coordination events, implement distinct features of neural population dynamics, and differentially contribute to the coordination of cross body movements that compose RGD. In particular, we uncovered a key role of cortico-thalamic communication in action ordering and coordination. Our findings reveal a broad concordance of cell type properties from molecular and anatomic features to neural population dynamics in a key area of a cortical network that articulates the action syntax of a complex behavior; they begin to integrate the neural circuit and population dynamics explanations of cortical function.

RESULTS

Target-guided reach and grasp to drink involves serial order and coordination of multiple forelimb and oral actions

The reach, grasp and withdraw-to-drink (RGD) task was performed in head-restrained mice (Figure 1A). In this task, mice use chemosensory and vibrissae cues to locate a waterspout in the dark with which to guide their left forelimb to grasp a water drop that they then withdraw to the mouth to drink by licking the hand [45]. Combining high-speed videography and deep neural network-based behavior tracking [38] (Figures S1A-H), we analyzed thirteen spatiotemporal profiles of the left hand and its relationship with the waterspout, mouth, and other body parts during RGD (Figure S1I). We thereby annotated four constituent movements: reach, grasp, withdraw, and hand lick (Figure 1B). The reach consists of a mouse *aiming* its hand after lifting it with the digits closed and flexed, *advancing* the hand with combined hand supination and digits extension, and *grasping* the target with digit flexing and closing (Figure 1C, Supplementary video 1). The reach endpoint at which a mouse grasped the water was accurate regardless of the targets changing locations (Figure 1D-E). After water grasping, the mouse withdraws the hand with a partial supination and brings it to the mouth with a further supination and finger opening. The mouth then opens and the tongue protrudes to lick and drink the water from the hand (Supplementary video 1); the hand remains supinated close to the mouth during

the multiple subsequent licks that consume the water (Figure 1F). These results indicate that RGD involves the orderly spatial and temporal coordination of multiple movements across the reach, grasp, withdraw and hand lick, such as digit opening during reach in anticipation of grasp, and supination toward the mouth during withdraw that is well-timed with tongue protrusion.

To examine the effect of variation of target location on the RGD movement, we presented the waterspout at five locations (Figure 1A): a central location (P3) aligned to the nose, two ipsilateral locations on the same side of reaching left hand (P1, P2), and two contralateral locations (P4, P5). After initial training with P2, mice retrieved water from each of 5 locations. In doing so, they used different forelimb trajectories that resulted from changing spatiotemporal coordination of the arm, wrist, and digit movements (Figure 1G-I, S1J-L). With changing target location, the mouse adjusts its aim by upper arm abduction or adduction and wrist flexion to point the digits to the target location (Figure 1G). This aiming phase during contralateral reaches occurs farther from the waterspout (Figure 1J) and requires larger angular corrections (Figure S1M) than ipsilateral reaches. The advance further adapts the palm-facing direction with digits opening and extension relative to target (Figure 1G), and both the aim and advance often involve adjusting movements revealed by changes of speed of the reaching hand (Figure 1G, S1N). These results indicate that reaching during RGD is not a ballistic movement but rather involves sensory feedback whereby mice adjust their arm, hand, and digit postures in relation to targets. Tongue protrusion before water grasping was rare, suggesting the orderly timing of forelimb and oral actions (Figure 1K, S1O).

Cortical network dynamics tracks RGD movement progression

Whereas the individual actions of reach, grasp, and lick can be elicited from spinal and brainstem motor centers [57,58], their orderly arrangement and coordination to compose a goal-directed RGD behavior likely involves a contribution from the cerebral cortex [59]. As RGD involves multiple constituent movements across the body as well as concurrent sensory inputs, its cortical control might include multiple sensory, motor, and other cortical areas. We used widefield calcium imaging to carry out an unbiased screen of neural activity dynamics across the dorsal cortex of mice performing RGD (Figure 2A, S2A, see Methods). By leveraging multiple mouse driver lines that target GCaMP6 expression to a set of molecular and projection defined PN types (Figure 2B), we monitored cortical network dynamics with PN type resolution [60,61]. Among these, the *Emx1-Cre* line targets most if not all PNs, *Cux1-CreER* targets L2/3 intratelencephalic (IT) neurons that project largely within the cortex, *PlxnD1-CreER* targets L2/3/5a IT subset with strong projection to cortex and striatum, *Fezf2-CreER* targets L5b pyramidal track (PT) and a small set of L6 output neurons, and *Tle4-CreER* targets L6 cortico-thalamic (CT) neurons (Figure 2B, S2B-C) [62]. These PN subpopulation driver lines allowed the detection of neural activity that is masked during whole-population recording (i.e. in the Emx1 line); driver line differences also provided physiological insight into the contribution of PN projection patterns.

Following training to reach to P2 with the left hand (Figure S2D), RGD was found to be accompanied by widespread activity dynamics in the contra- as well as ipsi-lateral hemisphere in all five PN populations (Figure S2E-F, Supplementary video 2). We then examined cortical network dynamics as

mice reached to the five different locations. Cortex-wide activity patterns in all five PN populations changed with RGD movements (Figure 2D) and were correlated with target location (Figure 2E). To summarize the network dynamics correlated with RGD progression to target locations, we built a generalized linear encoding model (GLM) considering ten forelimb movement profiles, including hand position and its relationship with target and mouth, to predict ”pixel-resolution” calcium fluorescent change (Figure S3A). With some regional differences, the model produced higher correlations for PN^{Emx1}, PT^{Fezf2}, and CT^{Tle4} populations than in IT^{Cux1} and IT^{PlxnD1} populations, as reflected by 10-fold cross-validated variance explained (R^2 value) (Figure 2F, S3B). A summary of GLM performance from all PNs revealed three activity nodes that were correlated with moment-by-moment forelimb movements: the central region of the secondary motor cortex (MOs-c, partially overlapping with rostral forelimb area (RFA) [21]), forelimb somatosensory area (SSp-ul), and parietal area (Prt) (Figure 2C-E, S3B-C). Analysis of IT^{Cux1} and IT^{PlxnD1} calcium signals additionally revealed nodes related to the anterior whisker barrel/unassigned sensory (SSp-bfd/un) and orofacial motor cortex (MO-orf) [63] (Figure S3B-C; see Methods for details of area definitions). Forelimb primary motor area (MOp-ul) [64] was isolated from GLM performance of IT^{Cux1}, corroborated by PT^{Fezf2} activity at the same area (Figure S3B). These results identify a cortical network with PN type resolution in which activity is correlated with the forelimb RGD movements.

As the target moved from P1 to P5, activity in the right hemisphere motor areas (contralateral to the left reaching hand) remained largely similar, while those in SSp-ul and SSp-bfd displayed decreased activity (Figure 2G), likely reflecting the shift of target locations. In parallel, there was an overall increase of activity in all ipsilateral areas except the parietal area (Figure 2G, Supplementary video 2). In sum, there was an overall relative increase of ipsilateral over contralateral activity as target moved from P1 to P5 (Figure 2G, Supplementary video 2). Indeed, activity of the network nodes decoded target location-related information (Figure S3D). Furthermore, within each area, the peak amplitude and temporal activity patterns of different PN populations were differentially modulated by target location (Figure S3E). The increased activation of IT^{Cux1} and IT^{PlxnD1} that occurred in ipsilateral sensory areas (SSp-bfd) as the waterspout moved from P1 to P5 likely reflects sensory information related to the changing target location (even though it was the left hand that was reaching). The increased activity of PT^{Fezf2} in left parietal and frontal motor areas as target location changed might reflect increased involvement of the right-side body to support more difficult contralateral reaches by the left forelimb. Together, the spatially-related hemispheric changes indicate that cortical network related to RGD includes the sensory and motor events related to the spatial positions of the target.

MOs-c is a necessary hub for the online articulation of RGD

To examine the functional role of each of the 6 network nodes (Figure 2C), we used a closed-loop optogenetic inhibition via light activation of Pvalb inhibitory interneurons in each of the six cortical areas in *Pvalb-ires-Cre;Ai32* mice [65] (Figure S4A). Inhibition was presented 129.2±42.8 msec after hand lift. Cortical inhibition contralateral to the reaching hand (cMOs-c) resulted in deficits of multiple RGD actions, including a decrease of grasp (measured by contact) and withdraw (measured by supination) (Figure 3A-B, S4B, Supplementary video 3). Inhibition of contralateral cMOp-ul resulted in a weaker but still significant decrease of supination probability (Figure 3B, S4B).

Inhibition of cSSp-ul resulted in deficits of digit closing, supination, and hand licking (Figure 3B, Supplementary video 4), suggesting that online somatosensory input may contribute to action progression and coordination. We did not observe significant effect upon cortical inhibition ipsilateral to the reaching hand (Figure S4B). In summary, among the six contralateral network nodes, MOs-c, MOp-ul, and SSp-ul seem essential for RGD movements.

We further analyzed when and how MOs-c suppression interfered RGD action progression in relation to each of the movements of reach, grasp, and withdraw to drink (Figure 3C). In about 1/5 of the trials, inhibition following hand lift resulted in deviation of subsequent hand movement away from target (Figure 3C, S4C-D). In trials when the hand did reach the target, supination during withdraw and/or hand lick often failed to occur (Figure 3C, S4E-F). Upon termination of inhibition, the mice immediately resumed and completed the RGD act (Figure S4E). To substantiate these results, we examined the effects of inhibiting MOs-c PNs directly by expressing a light sensitive inhibitory opsin GtACR1 [66] in *Emx1-Cre* mice using a Cre-dependent AAV vector (Figure 3D, S4G). Closed-loop inhibition of cMOs-c PNs resulted in similar impairments as those obtained with the activation of cMOs-c Pvalb interneurons (Figure S4H).

To further clarify the role of MOs-c for the RGD act, we gave prolonged inhibition spanning the entire trial, starting from one second before water delivery to the end of drinking. Prolonged inhibition of MOs-c PNs^{Emx1} only slightly decreased lift probability to more difficult contralateral targets (P4, P5) (Figure S4I), suggesting minor, if any, involvement of MOs-c in the initiation of RGD. On the other hand, consistent with closed-loop inhibition, prolonged inhibition of PNs^{Emx1} in contralateral or bilateral MOs-c decreased the probability of movement progression at multiple stages from reach, grasp, withdraw (measured as aim, contact, and supination, respectively) to hand lick (Figure 3E). Furthermore, cMOs-c but not iMOs-c PNs^{Emx1} inhibition resulted in a target location-dependent decrease in aim and grasp probability, with more severe deficits for more difficult contralateral targets (Figure 3F-G). In trials when the hand did reach the waterspout, we observed reduced digit opening at the reach endpoint in preparation of grasp (Figure 3H), attenuation of supinate during withdraw for drink (Figure 3I), and uncoordinated hand-oral movements for drinking (Figures 3J-O). Indeed, prolonged inhibition resulted in increased premature tongue protrusion before grasp (Figure 3L), abnormal hand posture (Figure 3M) and increased variation of hand position (Figure 3N) upon lick, and decreased correlation between hand and mouth movements during drink (Figure 3O). These deficits were not observed when inhibiting the iMOs-c PNs^{Emx1} (Figure S4J). Altogether, these results suggest that MOs-c is not crucial for the execution of specific individual actions (e.g. lift, advance, grasp, withdraw, and lick) but is involved in the orderly progression and coordination of these actions, especially the coordination between hand and oral actions.

MOs-c neural population activity correlates with RGD movements

To explore the neural coding properties and population dynamics within MOs-c, we performed electrophysiological recordings with linear probes during the RGD act (Figure 4A). Among a total of 1655 units across layers in 111 sessions from 6 mice, individual neurons were found to exhibit

diverse spiking patterns, with varying peak amplitude and latency relative to hand lift (Figure S5A-C). After hand lift, the activity of most activated neurons (86% of 921 units) peaked during reach, grasp and withdraw, only 10% peaked during the licking and drinking (Figure S5D-E). Deep layer neurons tended to be more strongly correlated with forelimb movements (Figure S5F-G), whereas upper layer neurons seemed more strongly modulated by target location (Figure S5H). Whereas some neurons (21.3%) exhibited a relatively brief increase of firing relative to baseline that correlated with more specific actions such as the reach or withdraw, most neurons (52%) exhibited more sustained spiking across adjacent actions, including transitions from one movement to the next, e.g. increasing (33%) or decreasing (19%) activity across reach and withdraw (Figure S5I).

To characterize the activity patterns of MOs-c neurons, we used a model-free computation approach [14] (see Methods) to group the activated neurons (921 units) into 7 clusters, each exhibiting distinct activity patterns in relation to movements and their targets locations (Figure 4B-D, S5J-K). Cluster 1 neuron activity corresponded to lift and decreased before the advance. Cluster 2 and 3 neuron activity peaked right after advance and then declined, but showed differential target modulation - either increased or decreased with more difficult contralateral targets. Cluster 4 neuron activity rose and peaked right after the advance and sustained their activity throughout withdraw and drinking, with little target modulation. Cluster 5 neurons were only mildly activated (187 units). Cluster 6 and 7 were neurons that were not significantly modulated (269 units) or inhibited (465 units) during RGD (Figure 4B-D). Overall, these results suggest that most MOs-c neurons collectively represented the RGD movement, with most neurons reflecting transitions across constituent movements.

To examine whether activity patterns of simultaneously recorded neurons encode moment-by-moment RGD movement progression, we built GLMs to predict the instantaneous hand position and other kinematic parameters within a session using population activity (see Methods). MOs-c population activity significantly decoded a range of arm, hand, and orofacial movement parameters (Figure 4E, S5L-M). The model predicted hand forward/upward movements and their relationship with the mouth and with the waterspout better than movement details such as path length, moving speed or digit open size (Figure 4E). In addition to movement parameters, MOs-c activity decoded target location, beginning even before hand lift, and with increasing accuracy during subsequent actions (Figure 4F, see Methods). These analyses suggest that MOs-c population activity correlates more closely with the relationships of arm movement to the target and with the mouth, than to individual arm or oral actions.

As cortical neural population dynamics have been implicated in forelimb motor control [67,68], we further investigated MOs-c population neural trajectories in high dimensional space during RGD (see Methods). We found that MOs-c population neural dynamics evolved with a smooth C-shaped trajectory in a low-dimension neural manifold along with the progression of RGD movements, turning upon hand lift, and concluding with another turn after advance endpoint before lick (Figure 4G, Supplementary video 5). Between different target locations, these trajectories shared the same geometric shape (not shown, all pairs of dissimilarity measurement Procrustes distances are close to zero) but progressively shifted in the neural manifold with target location (Figure 4G). This result suggests a robust neural substrate within MOs-c for generating a reliable pattern of population neural dynamics during RGD, which is systematically modulated by target location. Therefore, despite the

complex relationship between individual neuronal activity and RGD constituent actions, major temporal patterns of MOs population neural trajectories unfold/evolve along with RGD movements, suggesting their role in the orderly coordination of constituent actions that compose RGD.

Distinct coding properties and population neural trajectories of MOs-c PN types

Because conventional silicon probe recordings do not distinguish inhibitory and excitatory neurons reliably nor distinct PN types within excitatory neurons, the resulting data cannot be readily correlated to neural circuit operation. Leveraging our PN type driver lines, we next examined whether different molecular and projection defined PNs exhibit distinct neural coding properties and population dynamics associated with RGD. We used a well-established optical tagging method [69,70] to identify neurons of each of 4 PN types (Figure 5A, S6A-6B, see Methods), 22 IT^{Cux1}, 17 IT^{PlxnD1}, 63 PT^{Fezf2}, and 30 CT^{Tle4} neurons (Figure S6C-D). The depth distribution of these neurons matched their expected anatomical location, with deep layer PT^{Fezf2} and CT^{Tle4} neurons showing higher basal firing frequency (Figure S6E). Although each of these 4 PN types showed some degree of heterogeneity in activity patterns during RGD, they exhibited categorically distinct patterns. For example, the activity of many PT^{Fezf2} neurons rose before hand lift, peaked during reach, and declined immediately after reach endpoint (Figure 5B-D, S7A-C), while those of CT^{Tle4} typically rose just before reach endpoint, peaked at grasp-to-withdraw transition, and maintained elevated during withdraw and drinking (Figure 5B-D, S7A-C). Overall, higher fraction of PT^{Fezf2} and CT^{Tle4} neurons correlated better with movement progression than that of IT^{Cux1} or IT^{PlxnD1} neurons (Figure S7D). All 4 PN types show target tuning: IT^{Cux1} and IT^{PlxnD1} showed higher activity during ipsilateral trials, while PT^{Fezf2} activity increased during contralateral trials, and CT^{Tle4} neurons showed only weak tuning (Figure S7E). A majority of activated PT^{Fezf2} and CT^{Tle4} neurons can be assigned to different computationally identified activity clusters (Figure 5E), further indicating a physiological distinction between thalamic and corticofugal projecting populations.

To examine whether the 4 PN types manifest distinct population dynamics characteristics, we examined their averaged population neural trajectories during RGD (Supplementary video 5). PT^{Fezf2} exhibited a smooth C-shaped population neural trajectory that was systematically shifted according to target location (Figure 5F, S7G-I); these features closely resemble the neural trajectories of the whole MOs-c population (Figure 4I) but show a more clear "transitional bend" at the advance endpoint. Interestingly, CT^{Tle4} neurons also exhibited a smooth yet elongated C-shape population trajectory with a sharp transition at the advance endpoint, but was largely unmodulated by target location (Figure 5F, S7G-I). On the other hand, IT^{Cux1} and IT^{PlxnD1} showed jerky population trajectories that varied according to target location, and their loop-like geometric shapes showed less consistency among target locations (Figure 5F, S7G-I). These results suggest that different PN types give rise to highly distinct population dynamics and likely make unique contributions to the overall population neural trajectories of MOs-c. Given their categorically distinct axon projection patterns, these PN type-characteristic neural dynamics are likely separately conveyed to specific thalamic (CT^{Tle4}) and corticofugal (PT^{Fezf2}) target areas that may implement different aspects motor control and

coordination. These results extend the distinction of major PN types from their molecular, anatomic, and physiological features to population neural dynamics during behavior.

Distinct roles of MOs-c PN types in action sequence and coordination during RGD

To examine whether and how different PN types contribute to the execution and coordination of RGD, we optogenetically inhibited specific PN populations by virally expressing inhibitory opsin GtACR1 in *Cux1-*, *PlxnD1-*, *Fezf2-*, or *Tle4-CreER* driver lines, respectively (Figure 6A, S8A). We first used a closed-loop configuration by triggering inhibition following hand lift (Figure S8B). Whereas inhibition of IT^{Cux1} did not interfere action progression, consistent with the weak activation observed in widefield imaging (Figure S2E) and electrophysiology (Figure 5D). Inhibition of the other 3 PN types resulted in significant reduction in the success rate of subsequent actions such as target contact, supination and lick after grasp (Figure S8C). Among the 3 PN types, PT^{Fezf2} and CT^{Tle4} inhibition produced more and larger deficits in contact after lift than that of IT^{PlxnD1} (Figure S8C).

To substantiate these results, we then used a prolonged inhibition protocol throughout the RGD act (Figure 6A, as in Figure 3D) and analyzed the effect on the execution and coordination of all constituent actions (Figure 6B-F, Figure S8D-E). Again, IT^{Cux1} inhibition had no significant effect on RGD except a slight increase of digit opening for grasp (Figure 6B, 6D). IT^{PlxnD1} inhibition led to a mild increase of premature lick (Figure S8E). On the other hand, PT^{Fezf2} and CT^{Tle4} inhibition resulted in multiple deficits in action progression and coordination, including decreased grasp following reach, decreased supination following grasp, increased premature lick, and decreased hand-mouth coordination for drinking (Figure 6B-G, S8E). Whereas the deficits of PT^{Fezf2} inhibition were waterspout target location dependent (Figure 6C), i.e. more prominent in more difficult contralateral trials, those of CT^{Tle4} inhibition were more consistent and less modulated by waterspout target location (Figure 6C). These results are consistent with the activity patterns of PT^{Fezf2} and CT^{Tle4} during RGD, the formal being target dependent (i.e. increasing with contralateral P4 and P5) and the latter only weakly modulated by waterspout target location (Figure 5D). In particular, CT^{Tle4} inhibition led to more prominent deficit in hand-mouth coordination as reflected by increased variance in hand position and hand rotation at the time of tongue protrusion for drinking (Figure 6F-G). Together, these results suggest differential contributions by different PN types in the execution of constituent movements and highlight the key role of cortico-thalamic communication mediated by CT^{Tle4} in the orderly progression of actions, especially the coordination of hand and mouth movements that compose RGD.

DISCUSSION

Machine learning approaches enable studying the neural basis of serial order in behavior

A prerequisite for understanding the neural basis of serial order is to achieve accurate and quantitative analysis of its hierarchical organization as well as its spatiotemporal coordination of constituent movements, “the syntax of act”, that compose the goal-directed behavior. The mouse reach-to-drink (RDG) behavior is representative of the many consummatory skilled forelimb actions of rodents and

primates in that it entails the serial order of getting a hand to an allocentric target then getting the hand to an egocentric mouth target. Combining high-speed videography and machine learning-based pose estimation algorithms [38], we have achieved accurate and automated tracking and quantification of RGD at a fine resolution, including the complete repertoire of forelimb and oral actions (e.g. digit closure/extension, supination, lick), their orderly progression and coordination, target modulation, and trial-by-trial variability. These data reveal the "action syntax" of RGD that arranges a set of elemental actions to integrate the allocentric reach-grasp with egocentric withdraw-drink that compose a reach-to-drink ethological behavior. The automated and quantitative RGD ethogram provide a fine-grained reference frame to relate and interpret neural recording and the effects of functional manipulation, including temporally precise closed-loop feedback control.

A cortical network and key area for serial ordering and coordination of actions in RGD

Whereas elemental forelimb and oral actions (e.g. reach, grasp, lick) can be elicited from spinal and brainstem nuclei [71,72], their orderly arrangement towards goal-directed behaviors is likely orchestrated in higher centers such as the cortex [59], which features a comprehensive motor as well as sensory map across the body and the external world [73]. Most previous studies of cortical control of forelimb behaviors have largely (often exclusively) focused on the broadly defined primary forelimb and higher order motor areas [5,10,18-22,44,74], despite the recognition that multiple other areas (e.g. sensory, parietal) also play important roles [23,24 ,75]. Here, through an unbiased areal screen with PN type resolution, we have identified a dynamic cortical network that correlates with the sequential forelimb and orofacial movements that compose RGD. Compared with most previous wide-field imaging studies in the mouse [39,40,76-79], two design features of our study enabled mapping this network. First, the 4 driver lines that target major PN types, in addition to the pan-PN Emx1 line, allowed us to resolve PN-characteristic areal activity patterns that were masked in the broad population signal. Second, automated annotation of high-resolution RGD movement profiles enabled revealing the robust correlation between instantaneous behavioral and neural activity signals, thus the modeling approach to delineate the dynamic cortical network.

Importantly, closed-loop photo-inhibition of network nodes indicates that MOs-c, MOp-ul, and SSp-ul play significant and different roles in the execution and coordination of constituent actions of RGD. In particular, MOs-c inhibition resulted multiple deficits not only in the progression and coordination among forelimb actions (reach, grasp, supinate) but also between forelimb and mouth movements (withdraw to hand lick), suggesting it as a key hub in the orderly articulation of RGD.
MOs-c is located within the rostral forelimb area (RFA), a broadly defined premotor region in rats [80] and mice [21,22] implicated in the control of limb movement. Previous studies of RFA have mostly examined its role in the execution of individual reach or grasp movements, but not in the coordination of sequential actions across the body that compose an integrated behavior. MOs-c is densely connected with both primary sensory and motor areas [63,81], thus having access to target as well as body-part information. We suggest that MOs-c might be analogous to primate premotor and/or supplemental motor areas implicated in sensory-guided coordination of complex movement [16,82-84] and in tracking progress throughout a movement [19]. Together, these results uncover the dynamics

operation of an overarching PN type-resolved cortical network, likely centered around MOs-c, as a cortical substrate for the serial ordering and coordination of a complex motor behavior. Alternative interpretations of our results include that the MOs might be involved in the detailed execution of each constituent movement or may play a more general role in regulating RGD. While our current study is focused on the MOs-c, future efforts will more thoroughly examine the role of each network node and their interactions to discover how they collectively issue motor command, integrate sensory information, and coordinate cross body movements during RGD. As PN types are endowed with their inherent molecular, connectivity, and physiological features, these results further establish a top-level road map for exploring the underlying cellular and neural circuit mechanisms.

MOs-c neural activity decode movements and shape population dynamics during RGD

Our electrophysiology recordings revealed that, despite their diverse firing patterns, MOs-c neurons can be grouped into several clusters with activity modulations integrating across or marking the transition of adjacent actions that compose RGD. Substantiating these observations and leveraging high-resolution and automated annotation of RGD ethogram, our modeling reveals that MOs-c population activities encode the moment-by-moment relationships of arm movement with the allocentric waterspout target and with the egocentric mouth target, more than individual arm and oral actions, during RGD. Importantly, MOs-c population neural dynamics evolved with a smooth trajectory along with the progression of RGD constituent actions in an allocentric target location dependent manner. Similar neural trajectories in the forelimb primary motor cortex has been shown to drive a prehension task [68,85,86]. Our results suggest that MOs-c population activities encode the instantaneous movements across forelimb (arm, hand, digit) and orofacial (mouth, tongue, nose) body parts, especially the relationships between arm movements with the allocentric waterspout target and with the egocentric mouth target, and may contribute to the coordination of RGD constituent movements. Future studies will examine and compare the neural coding and population dynamics in other areas (especially MOp-ul and SSp-ul) and manipulate activity in sensory area while recording from motor areas (and vice versa); such studies will facilitate deciphering the dynamic processing across the RGD network that may underlie the behavior.

Cell type contribution to neural population dynamics and action coordination in RGD

Over the past two decades, studies of cortical control of movements have advanced beyond the historically influential “representational perspective” to exploring a “dynamical systems perspective” [29], which has emerged as a powerful candidate framework for understanding motor control and a range of brain functions [30-34,87]. However, a fundamental limitation of the current formulation of the dynamical systems framework is its missing link to the diverse cellular constituents and to neural circuits within which neural dynamics emerge, flow, and output [35]. To a large extent, this is because most electrophysiology studies provide little information about the genetic and anatomic identity of recorded neuron types, the basic elements of neural circuit. Here, we initiate a genetic dissection of motor cortex population dynamics by targeting 4 major PN types. Previous studies have highlighted their distinctions in transcription profiles [88], projection and local connectivity patterns [89], and intrinsic and synaptic properties [89,90]. We now demonstrate that they further manifest distinct features of

population dynamics during the execution and coordination of a complex motor behavior. Among these, the PT^{Fezf2} dynamics (Figure 5F) most closely resemble that of the MOs population (Figure 4J), suggesting it as a major component of the MOs neural dynamics. Interestingly, CT^{Tle4} activities also evolve with a robust neural trajectory that is largely not modulated by target location, suggesting a "generic" form of cortico-thalamic interactions during RGD. On the other hand, IT^{PlxnD1} and IT^{Cux1} show jerky neural trajectories without a prominent shape. Together, these results suggest that neural population dynamics in MOs may be composed from a set of elemental "dynamic themes" derived from neuronal subpopulations or cell types. It is possible that these cell-type dynamic themes might represent the "neural modes" or "latent variables" thought to constitute the building blocks of the population dynamics "symphony"[30].

As PN types are endowed with inherent physiological (intrinsic and synaptic) and anatomical (local and long-range connectivity) properties, in part shaped by their gene expression profiles [62,64,90,91], they represent physiological as well as structural building blocks of neural circuit dynamics; their inherent projection pattern further channel the outcome of cortical computation to cascades of brain areas [62,91] that drive behavior. For example, the intrinsic properties [90] and strong recurrent excitation [92-94] among L5B PT^{Fezf2} likely amplify the spiking signals of this circuit "motif", and their extensive corticofugal axon arbors undoubtedly broadcast their dynamic output across an extensive set of subcortical targets [62]. Importantly, targeted optogenetic inhibition further reveal that the 4 MOs PN types differentially contribute to the execution and coordination of RGD behavior. Whereas PT^{Fezf2} mostly contribute to the execution and coordination of forelimb actions especially when reaching for more difficult contralateral targets, CT^{Tle4} contribute to the orderly coordination of forelimb and oral actions less dependent on target locations. These results are consistent with their activity pattern and neural trajectories and highlight the crucial role of cortico-thalamic communications [95-97] in the coordination of cross body actions that compose RGD. Our findings thus reveal the broad multi-modal concordance of "cell type" properties from genetic, anatomical, and physiological characteristic to corresponding roles in neural population dynamics and behavioral function. They further demonstrate the feasibility of cell-type resolution analysis of population neural dynamics in mammalian cortex.

Altogether, our results delineate a cortical network and a key area and PN types therein that articular the serial order and coordination of a skilled behavior; they begin to integrate the neural circuit and neural manifold explanation of cortical computation that control behavior [98].

ACKNOWLEDGMENTS

We thank Drs. Nuo Li, Stephen Lisberger, Matthew Kaufman for helpful comments on the manuscript; Tatiana Engel, Yanliang Shi, William Galbavy, Shreyas M. Suryanarayana, Katherine S. Matho, and Dhananjay Huilgol for helpful discussions on the project; Ann Churchland and Simon Musall for advice on building wide-field imaging setup; Priscilla Wu, Joshua Hatfield, and Bao-Xia Han for animal preparation and maintenance; Weixin Zhong for schematic drawings. This work was supported by NIMH research project U19MH114823-01 grant to Z.J.H. Z.J.H. is supported by an NIH Director's Pioneer Award 1DP1MH129954-01.

AUTHOR CONTRIBUTIONS

Y.L. and Z.J.H. conceived the study. Z.J.H. acquired the funding, managed the project, and supervised the study. Y.L. designed experiments, built setups, performed experiments, established pipelines, analyzed the data, and presented for visualization. X.A. provided animals, helped building rigs, and shared materials and resources. Y.Q. prepared AAV vectors and helped with *in vivo* electrophysiological recording. X.H.X contributed to behavioral training and optogenetic manipulation. S.Z. packaged AAV vectors and carried out *in situ* hybridization. H.M. helped with the wide-field imaging rig. L.B.R. and N.B. helped analyzing and interpreting *in vivo* electrophysiological data. I.Q.W. set principles for behavioral annotations. Z.J.H. and Y.L. wrote the manuscript with significant edits from I.Q.W. All authors discussed the data, interpreted the data, offered advice for data analysis, reviewed, and provided inputs to the manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interests.

INCLUSION AND DIVERSITY

We support inclusive, diverse, and equitable conduct of research.

RESOURCE AVAILABILITY

Animals and materials availability

All mice, reagents and materials are openly available with detail source information listed in the Methods.

Data and code availability

All original data reported in this study and any additional information required to reanalyze the data reported in this paper will be shared by correspondence authors upon request. Customized code related to data analysis will be available on GitHub (https://github.com/liyistart/RGD-2023).

REFERENCES

1. Iwaniuk, A.N., and Whishaw, I.Q. (2000). On the origin of skilled forelimb movements. Trends in Neurosciences.
2. Whishaw, I.Q., and Karl, J.M. (2019). The evolution of the hand as a tool in feeding behavior: The multiple motor channel theory of hand use. In Feeding in vertebrates: Evolution, morphology, behavior, biomechanics, V. Bels, and I.Q. Whishaw, eds. (Springer International Publishing), pp. 159-186. 10.1007/978-3-030-13739-7_6.
3. Lashley, K.S. (1951). The problem of serial order in behavior. In Cerebral mechanisms in behavior; the hixon symposium., (Wiley), pp. 112-146.
4. Rosenbaum, R.S., Stuss, D.T., Levine, B., and Tulving, E. (2007). Theory of mind is independent of episodic memory. Science *318*, 1257. 10.1126/science.1148763.
5. Evarts, E.V. (1968). Relation of pyramidal tract activity to force exerted during voluntary movement. J Neurophysiol *31*, 14-27. 10.1152/jn.1968.31.1.14.
6. Georgopoulos, A., Kalaska, J., Caminiti, R., and Massey, J. (1982). On the relations between the direction of two-dimensional arm movements and cell discharge in primate motor cortex. The Journal of Neuroscience *2*, 1527-1537. 10.1523/jneurosci.02-11-01527.1982.
7. Humphrey, D.R., Schmidt, E.M., and Thompson, W.D. (1970). Predicting measures of motor performance from multiple cortical spike trains. Science *170*, 758-762. 10.1126/science.170.3959.758.
8. Morrow, M.M., Jordan, L.R., and Miller, L.E. (2007). Direct comparison of the task-dependent discharge of m1 in hand space and muscle space. J Neurophysiol *97*, 1786-1798. 10.1152/jn.00150.2006.
9. Thach, W.T. (1978). Correlation of neural discharge with pattern and force of muscular activity, joint position, and direction of intended next movement in motor cortex and cerebellum. J Neurophysiol *41*, 654-676. 10.1152/jn.1978.41.3.654.
10. Mushiake, H., Inase, M., and Tanji, J. (1991). Neuronal activity in the primate premotor, supplementary, and precentral motor cortex during visually guided and internally determined sequential movements. J Neurophysiol *66*, 705-718. 10.1152/jn.1991.66.3.705.
11. Nakamura, K., Sakai, K., and Hikosaka, O. (1998). Neuronal activity in medial frontal cortex during learning of sequential procedures. J Neurophysiol *80*, 2671-2687. 10.1152/jn.1998.80.5.2671.
12. Shima, K., and Tanji, J. (1998). Both supplementary and presupplementary motor areas are crucial for the temporal organization of multiple movements. J Neurophysiol *80*, 3247-3260. 10.1152/jn.1998.80.6.3247.
13. Tanji, J. (2001). Sequential organization of multiple movements: Involvement of cortical motor areas. Annual Review of Neuroscience *24*, 631-651. 10.1146/annurev.neuro.24.1.631.
14. Xu, D., Dong, M., Chen, Y., Delgado, A.M., Hughes, N.C., Zhang, L., and O'Connor, D.H. (2022). Cortical processing of flexible and context-dependent sensorimotor sequences. Nature *603*, 464-469. 10.1038/s41586-022-04478-7.
15. Wong, A.L., and Krakauer, J.W. (2019). Why are sequence representations in primary motor cortex so elusive? Neuron *103*, 956-958. 10.1016/j.neuron.2019.09.011.
16. Graziano, M.S.A. (2016). Ethological action maps: A paradigm shift for the motor cortex. Trends Cogn Sci *20*, 121-132. 10.1016/j.tics.2015.10.008.
17. Penfield, W., and Welch, K. (1951). The supplementary motor area of the cerebral cortex: A clinical and experimental study. Archives of Neurology And Psychiatry *66*, 289-317. 10.1001/archneurpsyc.1951.02320090038004.
18. Kaufman, M.T., Churchland, M.M., Santhanam, G., Yu, B.M., Afshar, A., Ryu, S.I., and Shenoy, K.V. (2010). Roles of monkey premotor neuron classes in movement preparation and execution. J Neurophysiol *104*, 799-810. 10.1152/jn.00231.2009.
19. Russo, A.A., Khajeh, R., Bittner, S.R., Perkins, S.M., Cunningham, J.P., Abbott, L.F., and Churchland, M.M. (2020). Neural trajectories in the supplementary motor area and motor cortex exhibit distinct geometries, compatible with different classes of computation. Neuron *107*, 745-758 e746. 10.1016/j.neuron.2020.05.020.

20. Strick, P.L., Dum, R.P., and Rathelot, J.A. (2021). The cortical motor areas and the emergence of motor skills: A neuroanatomical perspective. Annu Rev Neurosci *44*, 425-447. 10.1146/annurev-neuro-070918-050216.
21. Tennant, K.A., Adkins, D.L., Donlan, N.A., Asay, A.L., Thomas, N., Kleim, J.A., and Jones, T.A. (2011). The organization of the forelimb representation of the c57bl/6 mouse motor cortex as defined by intracortical microstimulation and cytoarchitecture. Cereb Cortex *21*, 865-876. 10.1093/cercor/bhq159.
22. Wang, X., Liu, Y., Li, X., Zhang, Z., Yang, H., Zhang, Y., Williams, P.R., Alwahab, N.S.A., Kapur, K., Yu, B., et al. (2017). Deconstruction of corticospinal circuits for goal-directed motor skills. Cell *171*, 440-455.e414. 10.1016/j.cell.2017.08.014.
23. Evarts, E.V. (1972). Contrasts between activity of precentral and postcentral neurons of cerebral cortex during movement in the monkey. Brain Res *40*, 25-31. 10.1016/0006-8993(72)90101-1.
24. Scott, S.H., Cluff, T., Lowrey, C.R., and Takei, T. (2015). Feedback control during voluntary motor actions. Current Opinion in Neurobiology. Elsevier Ltd.
25. Hatsopoulos, N.G., and Donoghue, J.P. (2009). The science of neural interface systems. Annu Rev Neurosci *32*, 249-266. 10.1146/annurev.neuro.051508.135241.
26. Kalaska, J.F. (2009). From intention to action: Motor cortex and the control of reaching movements. Adv Exp Med Biol *629*, 139-178. 10.1007/978-0-387-77064-2_8.
27. Fetz, E.E. (1992). Are movement parameters recognizably coded in the activity of single neurons? Behavioral and Brain Sciences *15*, 679-690.
28. Scott, S.H. (2008). Inconvenient truths about neural processing in primary motor cortex. J Physiol *586*, 1217-1224. 10.1113/jphysiol.2007.146068.
29. Shenoy, K.V., Sahani, M., and Churchland, M.M. (2013). Cortical control of arm movements: A dynamical systems perspective. Annual Review of Neuroscience *36*, 337-359. 10.1146/annurev-neuro-062111-150509.
30. Gallego, J.A., Perich, M.G., Miller, L.E., and Solla, S.A. (2017). Neural manifolds for the control of movement. Neuron. Cell Press.
31. Bernardi, S., Benna, M.K., Rigotti, M., Munuera, J., Fusi, S., and Salzman, C.D. (2020). The geometry of abstraction in the hippocampus and prefrontal cortex. Cell *183*, 954-967 e921. 10.1016/j.cell.2020.09.031.
32. Chung, S., and Abbott, L.F. (2021). Neural population geometry: An approach for understanding biological and artificial neural networks. Curr Opin Neurobiol *70*, 137-144. 10.1016/j.conb.2021.10.010.
33. Azeredo da Silveira, R., and Rieke, F. (2021). The geometry of information coding in correlated neural populations. Annu Rev Neurosci *44*, 403-424. 10.1146/annurev-neuro-120320-082744.
34. Vyas, S., Golub, M.D., Sussillo, D., and Shenoy, K.V. (2020). Computation through neural population dynamics. Annual Review of Neuroscience *43*, 249-275. 10.1146/annurev-neuro-092619-094115.
35. O'Shea, D.J., Trautmann, E., Chandrasekaran, C., Stavisky, S., Kao, J.C., Sahani, M., Ryu, S., Deisseroth, K., and Shenoy, K.V. (2017). The need for calcium imaging in nonhuman primates: New motor neuroscience and brain-machine interfaces. Exp Neurol *287*, 437-451. 10.1016/j.expneurol.2016.08.003.
36. An, X., Matho, K., Li, Y., Mohan, H., Xu, X.H., Whishaw, I.Q., Kepecs, A., and Huang, Z.J. (2022). A cortical circuit for orchestrating oromanual food manipulation. bioRxiv, 2022.2012.2003.518964. 10.1101/2022.12.03.518964.
37. Barrett, J.M., Raineri Tapies, M.G., and Shepherd, G.M.G. (2020). Manual dexterity of mice during food-handling involves the thumb and a set of fast basic movements. PLoS One *15*, e0226774. 10.1371/journal.pone.0226774.
38. Mathis, A., Mamidanna, P., Cury, K.M., Abe, T., Murthy, V.N., Mathis, M.W., and Bethge, M. (2018). Deeplabcut: Markerless pose estimation of user-defined body parts with deep learning. Nature Neuroscience *21*, 1281-1289. 10.1038/s41593-018-0209-y.

39. Allen, W.E., Kauvar, I.V., Chen, M.Z., Richman, E.B., Yang, S.J., Chan, K., Gradinaru, V., Deverman, B.E., Luo, L., and Deisseroth, K. (2017). Global representations of goal-directed behavior in distinct cell types of mouse neocortex. Neuron *94*, 891-907.e896. 10.1016/j.neuron.2017.04.017.
40. Makino, H., Ren, C., Liu, H., Kim, A.N., Kondapaneni, N., Liu, X., Kuzum, D., and Komiyama, T. (2017). Transformation of cortex-wide emergent properties during motor learning. Neuron *94*, 880-890 e888. 10.1016/j.neuron.2017.04.015.
41. Musall, S., Kaufman, M.T., Juavinett, A.L., Gluf, S., and Churchland, A.K. (2019). Single-trial neural dynamics are dominated by richly varied movements. Nat Neurosci *22*, 1677-1686. 10.1038/s41593-019-0502-4.
42. Hausmann, S.B., Vargas, A.M., Mathis, A., and Mathis, M.W. (2021). Measuring and modeling the motor system with machine learning. Curr Opin Neurobiol *70*, 11-23. 10.1016/j.conb.2021.04.004.
43. Wagner, M.J., Savall, J., Hernandez, O., Mel, G., Inan, H., Rumyantsev, O., Lecoq, J., Kim, T.H., Li, J.Z., Ramakrishnan, C., et al. (2021). A neural circuit state change underlying skilled movements. Cell *184*, 3731-3747 e3721. 10.1016/j.cell.2021.06.001.
44. Guo, J.Z., Graves, A.R., Guo, W.W., Zheng, J., Lee, A., Rodríguez-Gonzá Lez, J., Li, N., Macklin, J.J., Phillips, J.W., Mensh, B.D., et al. (2015). Cortex commands the performance of skilled movement. eLife *4*. 10.7554/eLife.10774.
45. Galiñanes, G.L., Bonardi, C., and Huber, D. (2018). Directional reaching for water as a cortex-dependent behavioral framework for mice. Cell Reports *22*, 2767-2783. 10.1016/j.celrep.2018.02.042.
46. Currie, S.P., Ammer, J.J., Premchand, B., Dacre, J., Wu, Y., Eleftheriou, C., Colligan, M., Clarke, T., Mitchell, L., Faisal, A.A., et al. (2022). Movement-specific signaling is differentially distributed across motor cortex layer 5 projection neuron classes. Cell Reports *39*, 110801. https://doi.org/10.1016/j.celrep.2022.110801.
47. Park, J., Phillips, J.W., Guo, J.-Z., Martin, K.A., Hantman, A.W., and Dudman, J.T. (2022). Motor cortical output for skilled forelimb movement is selectively distributed across projection neuron classes. Science Advances *8*, eabj5167. doi:10.1126/sciadv.abj5167.
48. Mosberger, A., Sibener, L., Chen, T., Rodrigues, H., Hormigo, R., Ingram, J., Athalye, V., Tabachnik, T., Wolpert, D., Murray, J., and Costa, R. (2023). Exploration biases how forelimb reaches to a spatial target are learned. bioRxiv, 2023.2005.2008.539291. 10.1101/2023.05.08.539291.
49. Takatoh, J., Prevosto, V., Thompson, P.M., Lu, J., Chung, L., Harrahill, A., Li, S., Zhao, S., He, Z., Golomb, D., et al. (2022). The whisking oscillator circuit. Nature *609*, 560-568. 10.1038/s41586-022-05144-8.
50. Mercer Lindsay, N., Knutsen, P.M., Lozada, A.F., Gibbs, D., Karten, H.J., and Kleinfeld, D. (2019). Orofacial movements involve parallel corticobulbar projections from motor cortex to trigeminal premotor nuclei. Neuron *104*, 765-780.e763. https://doi.org/10.1016/j.neuron.2019.08.032.
51. Petersen, C.C.H. (2019). Sensorimotor processing in the rodent barrel cortex. Nat Rev Neurosci *20*, 533-546. 10.1038/s41583-019-0200-y.
52. Li, N., Chen, T.W., Guo, Z.V., Gerfen, C.R., and Svoboda, K. (2015). A motor cortex circuit for motor planning and movement. Nature *519*, 51-56. 10.1038/nature14178.
53. Economo, M.N., Viswanathan, S., Tasic, B., Bas, E., Winnubst, J., Menon, V., Graybuck, L.T., Nguyen, T.N., Smith, K.A., Yao, Z., et al. (2018). Distinct descending motor cortex pathways and their roles in movement. Nature *563*, 79-84. 10.1038/s41586-018-0642-9.
54. Gao, Z., Davis, C., Thomas, A.M., Economo, M.N., Abrego, A.M., Svoboda, K., De Zeeuw, C.I., and Li, N. (2018). A cortico-cerebellar loop for motor planning. Nature *563*, 113-116. 10.1038/s41586-018-0633-x.
55. Bollu, T., Ito, B.S., Whitehead, S.C., Kardon, B., Redd, J., Liu, M.H., and Goldberg, J.H. (2021). Cortex-dependent corrections as the tongue reaches for and misses targets. Nature, 1-6. 10.1038/s41586-021-03561-9.
56. Barrett, J.M., Martin, M.E., and Shepherd, G.M.G. (2022). Manipulation-specific cortical activity as mice handle food. Current Biology *32*, 4842-4853.e4846. https://doi.org/10.1016/j.cub.2022.09.045.

57. Ruder, L., Schina, R., Kanodia, H., Valencia-Garcia, S., Pivetta, C., and Arber, S. (2021). A functional map for diverse forelimb actions within brainstem circuitry. Nature *590*, 445-450. 10.1038/s41586-020-03080-z.
58. Arber, S., and Costa, R.M. (2018). Connecting neuronal circuits for movement. Science. American Association for the Advancement of Science.
59. Merel, J., Botvinick, M., and Wayne, G. (2019). Hierarchical motor control in mammals and machines. Nature Communications *10*, 5489. 10.1038/s41467-019-13239-6.
60. Mohan, H., An, X., Xu, X.H., Kondo, H., Zhao, S., Matho, K.S., Wang, B.S., Musall, S., Mitra, P., and Huang, Z.J. (2023). Cortical glutamatergic projection neuron types contribute to distinct functional subnetworks. Nat Neurosci *26*, 481-494. 10.1038/s41593-022-01244-w.
61. Musall, S., Sun, X.R., Mohan, H., An, X., Gluf, S., Li, S.J., Drewes, R., Cravo, E., Lenzi, I., Yin, C., et al. (2023). Pyramidal cell types drive functionally distinct cortical activity patterns during decision-making. Nat Neurosci *26*, 495-505. 10.1038/s41593-022-01245-9.
62. Matho, K.S., Huilgol, D., Galbavy, W., He, M., Kim, G., An, X., Lu, J., Wu, P., Di Bella, D.J., Shetty, A.S., et al. (2021). Genetic dissection of the glutamatergic neuron system in cerebral cortex. Nature *598*, 182-187. 10.1038/s41586-021-03955-9.
63. Zingg, B., Hintiryan, H., Gou, L., Song, M.Y., Bay, M., Bienkowski, M.S., Foster, N.N., Yamashita, S., Bowman, I., Toga, A.W., and Dong, H.W. (2014). Neural networks of the mouse neocortex. Cell *156*, 1096-1111. 10.1016/j.cell.2014.02.023.
64. Callaway, E.M., Dong, H.-W., Ecker, J.R., Hawrylycz, M.J., Huang, Z.J., Lein, E.S., Ngai, J., Osten, P., Ren, B., Tolias, A.S., et al. (2021). A multimodal cell census and atlas of the mammalian primary motor cortex. Nature *598*, 86-102. 10.1038/s41586-021-03950-0.
65. Li, N., Chen, S., Guo, Z.V., Chen, H., Huo, Y., Inagaki, H.K., Chen, G., Davis, C., Hansel, D., Guo, C., and Svoboda, K. (2019). Spatiotemporal constraints on optogenetic inactivation in cortical circuits. eLife *8*. 10.7554/eLife.48622.
66. Govorunova, E.G., Sineshchekov, O.A., Janz, R., Liu, X., and Spudich, J.L. (2015). Natural light-gated anion channels: A family of microbial rhodopsins for advanced optogenetics. Science *349*, 647-650. 10.1126/science.aaa7484.
67. Churchland, M.M., Cunningham, J.P., Kaufman, M.T., Foster, J.D., Nuyujukian, P., Ryu, S.I., Shenoy, K.V., and Shenoy, K.V. (2012). Neural population dynamics during reaching. Nature *487*, 51-56. 10.1038/nature11129.
68. Sauerbrei, B.A., Guo, J.Z., Cohen, J.D., Mischiati, M., Guo, W., Kabra, M., Verma, N., Mensh, B., Branson, K., and Hantman, A.W. (2020). Cortical pattern generation during dexterous movement is input-driven. Nature *577*, 386-391. 10.1038/s41586-019-1869-9.
69. Kvitsiani, D., Ranade, S., Hangya, B., Taniguchi, H., Huang, J.Z., and Kepecs, A. (2013). Distinct behavioural and network correlates of two interneuron types in prefrontal cortex. Nature *498*, 363-366. 10.1038/nature12176.
70. Li, Y., Zhong, W., Wang, D., Feng, Q., Liu, Z., Zhou, J., Jia, C., Hu, F., Zeng, J., Guo, Q., et al. (2016). Serotonin neurons in the dorsal raphe nucleus encode reward signals. Nature Communications *7*, 10503. 10.1038/ncomms10503.
71. Arber, S., and Costa, R.M. (2022). Networking brainstem and basal ganglia circuits for movement. Nat Rev Neurosci *23*, 342-360. 10.1038/s41583-022-00581-w.
72. Travers, J.B., Dinardo, L.A., and Karimnamazi, H. (1997). Motor and premotor mechanisms of licking. Neurosci Biobehav Rev *21*, 631-647. 10.1016/s0149-7634(96)00045-0.
73. Douglas, R.J., and Martin, K.A. (2012). Behavioral architecture of the cortical sheet. Curr Biol *22*, R1033-1038. 10.1016/j.cub.2012.11.017.
74. Yang, W., Kanodia, H., and Arber, S. (2023). Structural and functional map for forelimb movement phases between cortex and medulla. Cell *186*, 162-177.e118. 10.1016/j.cell.2022.12.009.
75. Gharbawie, O.A., Stepniewska, I., Burish, M.J., and Kaas, J.H. (2010). Thalamocortical connections of functional zones in posterior parietal cortex and frontal cortex motor regions in new world monkeys. Cereb Cortex *20*, 2391-2410. 10.1093/cercor/bhp308.

76. Wekselblatt, J.B., Flister, E.D., Piscopo, D.M., and Niell, C.M. (2016). Large-scale imaging of cortical dynamics during sensory perception and behavior. J Neurophysiol *115*, 2852-2866. 10.1152/jn.01056.2015.
77. Vanni, M.P., Chan, A.W., Balbi, M., Silasi, G., and Murphy, T.H. (2017). Mesoscale mapping of mouse cortex reveals frequency-dependent cycling between distinct macroscale functional modules. Journal of Neuroscience *37*, 7513-7533. 10.1523/JNEUROSCI.3560-16.2017.
78. Quarta, E., Scaglione, A., Lucchesi, J., Sacconi, L., Allegra Mascaro, A.L., and Pavone, F.S. (2022). Distributed and localized dynamics emerge in the mouse neocortex during reach-to-grasp behavior. The Journal of Neuroscience *42*, 777-788. 10.1523/jneurosci.0762-20.2021.
79. Wang, Y., Sepers, M.D., Xiao, D., Raymond, L.A., and Murphy, T.H. (2023). Water-reaching platform for longitudinal assessment of cortical activity and fine motor coordination defects in a huntington disease mouse model. eNeuro. 10.1523/ENEURO.0452-22.2022.
80. Rouiller, E.M., Moret, V., and Liang, F. (1993). Comparison of the connectional properties of the two forelimb areas of the rat sensorimotor cortex: Support for the presence of a premotor or supplementary motor cortical area. Somatosens Mot Res *10*, 269-289. 10.3109/08990229309028837.
81. Gamanut, R., Kennedy, H., Toroczkai, Z., Ercsey-Ravasz, M., Van Essen, D.C., Knoblauch, K., and Burkhalter, A. (2018). The mouse cortical connectome, characterized by an ultra-dense cortical graph, maintains specificity by distinct connectivity profiles. Neuron *97*, 698-715 e610. 10.1016/j.neuron.2017.12.037.
82. Wise, S.P. (1985). The primate premotor cortex: Past, present, and preparatory. Annual Review of Neuroscience *8*, 1-19. 10.1146/annurev.ne.08.030185.000245.
83. Hoshi, E., and Tanji, J. (2000). Integration of target and body-part information in the premotor cortex when planning action. Nature *408*, 466-470. 10.1038/35044075.
84. Zimnik, A.J., and Churchland, M.M. (2021). Independent generation of sequence elements by motor cortex. Nat Neurosci *24*, 412-424. 10.1038/s41593-021-00798-5.
85. Levy, S., Lavzin, M., Benisty, H., Ghanayim, A., Dubin, U., Achvat, S., Brosh, Z., Aeed, F., Mensh, B.D., Schiller, Y., et al. (2020). Cell-type-specific outcome representation in the primary motor cortex. Neuron. 10.1016/j.neuron.2020.06.006.
86. Zhu, F., Grier, H.A., Tandon, R., Cai, C., Agarwal, A., Giovannucci, A., Kaufman, M.T., and Pandarinath, C. (2022). A deep learning framework for inference of single-trial neural population dynamics from calcium imaging with subframe temporal resolution. Nat Neurosci *25*, 1724-1734. 10.1038/s41593-022-01189-0.
87. Barack, D.L., and Krakauer, J.W. (2021). Two views on the cognitive brain. Nat Rev Neurosci *22*, 359-371. 10.1038/s41583-021-00448-6.
88. Yao, Z., van Velthoven, C.T.J., Nguyen, T.N., Goldy, J., Sedeno-Cortes, A.E., Baftizadeh, F., Bertagnolli, D., Casper, T., Chiang, M., Crichton, K., et al. (2021). A taxonomy of transcriptomic cell types across the isocortex and hippocampal formation. Cell *184*, 3222-3241 e3226. 10.1016/j.cell.2021.04.021.
89. Harris, K.D., and Shepherd, G.M.G. (2015). The neocortical circuit: Themes and variations. Nature Neuroscience. Nature Publishing Group.
90. Scala, F., Kobak, D., Bernabucci, M., Bernaerts, Y., Cadwell, C.R., Castro, J.R., Hartmanis, L., Jiang, X., Laturnus, S., Miranda, E., et al. (2021). Phenotypic variation of transcriptomic cell types in mouse motor cortex. Nature *598*, 144-150. 10.1038/s41586-020-2907-3.
91. Munoz-Castaneda, R., Zingg, B., Matho, K.S., Chen, X., Wang, Q., Foster, N.N., Li, A., Narasimhan, A., Hirokawa, K.E., Huo, B., et al. (2021). Cellular anatomy of the mouse primary motor cortex. Nature *598*, 159-166. 10.1038/s41586-021-03970-w.
92. Le Be, J.V., Silberberg, G., Wang, Y., and Markram, H. (2007). Morphological, electrophysiological, and synaptic properties of corticocallosal pyramidal cells in the neonatal rat neocortex. Cereb Cortex *17*, 2204-2213. 10.1093/cercor/bhl127.
93. Shepherd, G.M.G. (2013). Corticostriatal connectivity and its role in disease. Nature Reviews Neuroscience. Nature Publishing Group.

94. Kawaguchi, Y. (2017). Pyramidal cell subtypes and their synaptic connections in layer 5 of rat frontal cortex. Cereb Cortex *27*, 5755-5771. 10.1093/cercor/bhx252.
95. Moll, F.W., Kranz, D., Corredera Asensio, A., Elmaleh, M., Ackert-Smith, L.A., and Long, M.A. (2023). Thalamus drives vocal onsets in the zebra finch courtship song. Nature *616*, 132-136. 10.1038/s41586-023-05818-x.
96. Shepherd, G.M.G., and Yamawaki, N. (2021). Untangling the cortico-thalamo-cortical loop: Cellular pieces of a knotty circuit puzzle. Nat Rev Neurosci *22*, 389-406. 10.1038/s41583-021-00459-3.
97. Carmona, L.M., Nelson, A., Tun, L.T., Kim, A., Shiao, R., Kissner, M.D., Menon, V., and Costa, R.M. (2023). Corticothalamic neurons in motor cortex have a permissive role in motor execution. bioRxiv, 2022.2009.2020.508799. 10.1101/2022.09.20.508799.
98. Langdon, C., Genkin, M., and Engel, T.A. (2023). A unifying perspective on neural manifolds and circuits for cognition. Nat Rev Neurosci *24*, 363-377. 10.1038/s41583-023-00693-x.
99. Guo, Z.V., Hires, S.A., Li, N., O'Connor, D.H., Komiyama, T., Ophir, E., Huber, D., Bonardi, C., Morandell, K., Gutnisky, D., et al. (2014). Procedures for behavioral experiments in head-fixed mice. PLoS One *9*, e88678. 10.1371/journal.pone.0088678.
100. Couto, J., Musall, S., Sun, X.R., Khanal, A., Gluf, S., Saxena, S., Kinsella, I., Abe, T., Cunningham, J.P., Paninski, L., and Churchland, A.K. (2021). Chronic, cortex-wide imaging of specific cell populations during behavior. Nat Protoc *16*, 3241-3263. 10.1038/s41596-021-00527-z.
101. Kane, G.A., Lopes, G., Saunders, J.L., Mathis, A., and Mathis, M.W. (2020). Real-time, low-latency closed-loop feedback using markerless posture tracking. Elife *9*. 10.7554/eLife.61909.
102. Whishaw, I.Q., and Tomie, J.A. (1989). Olfaction directs skilled forelimb reaching in the rat. Behav Brain Res *32*, 11-21. 10.1016/s0166-4328(89)80067-1.
103. Whishaw, I.Q., and Pellis, S.M. (1990). The structure of skilled forelimb reaching in the rat: A proximally driven movement with a single distal rotatory component. Behavioural Brain Research *41*, 49-59. 10.1016/0166-4328(90)90053-H.
104. Saxena, S., Kinsella, I., Musall, S., Kim, S.H., Meszaros, J., Thibodeaux, D.N., Kim, C., Cunningham, J., Hillman, E.M.C., Churchland, A., and Paninski, L. (2020). Localized semi-nonnegative matrix factorization (locanmf) of widefield calcium imaging data. PLoS Comput Biol *16*, e1007791. 10.1371/journal.pcbi.1007791.
105. Vaccari, F.E., Diomedi, S., Filippini, M., Galletti, C., and Fattori, P. (2021). A poisson generalized linear model application to disentangle the effects of various parameters on neurophysiological discharges. STAR Protoc *2*, 100413. 10.1016/j.xpro.2021.100413.
106. Gallego, J.A., Perich, M.G., Chowdhury, R.H., Solla, S.A., and Miller, L.E. (2020). Long-term stability of cortical population dynamics underlying consistent behavior. Nat Neurosci *23*, 260-270. 10.1038/s41593-019-0555-4.

FIGURE LEGEND

Figure 1. Orderly coordination of constituent movements during reach and grasp to drink (RGD).

(A) Schematic of a head-restrained mouse reaching with its left hand for a waterspout positioned randomly at one of five locations (ipsilateral P1, P2; central P3; contralateral P4, P5).

(B) Schematic of the serial order of RGD and its constituent movements. RGD involves four components: reach, grasp, withdraw and hand lick. The reach consists of hand lifting, aiming, and advancing to target; the grasp opens and closes the digits at the target; the withdraw supinates the hand to the mouth and opens the digits to enable water licking. The color code is used in all subsequent figures.

(C) Example hand trajectory annotated with constituent actions. The dashed line indicates the reference direction in relation to target; +, waterspout location.

(D) Spatial contour map of hand location from side view at lift, aim and advance endpoint. Contours indicate probability starting from $0.01/mm^2$ with equal increment of $0.01/mm^2$; 6229 trials from 70 sessions in 25 mice across five target locations; +, waterspout location.

(E) Hand position relative to target at the advance endpoint (median ± interquartile range). Zero means digits are located at the waterspout. Kruskal-Wallis test with Tukey's post hoc test, χ^2 = 40.74, p = 1.18e-05. Horizontal lines in boxplots indicate 75%, 50%, and 25% percentile. Whiskers represent data point span to 90% or 10% percentile. Data from 70 sessions from 25 mice. The same animals were used in the following figure panels.

(F) Top: Hand postures at lift (fingers slightly closed and flexed), aim (palm rotated toward target), advance endpoint (fingers extend and open for grasp) and hand-lick (hand can be closed or open). Bottom: Distribution of hand rotation score as a reflection of palm-facing direction. Data from 6229 lifts, aims, advance endpoints; and 119584 licks from 70 sessions in 25 mice. Note the near 180-degree hand supination from reach-grasp (pronated) to withdraw-lick (supinated).

(G) Movement profiles of an example trial from lift onset to first hand lick. Annotated lift, aim, advance, grasp/withdraw are color coded. Dashed lines indicate z-score normalized zeros for each movement variable. Arrows point to two separate hand speed accelerations that reflect movement adjustments during reaching.

(H) Example of annotated hand trajectories of three random trials for each position in the same session.

(I) Movement duration from lift to grasp for the five target locations. Two-tailed nonparametric Kruskal-Wallis test, statistic χ^2 = 91.07, p = 7.78e-19.

(J) Distance from aim to target for the five locations. Kruskal-Wallis test, χ^2 = 100.02, p = 9.72e-21.

(K) Distributions of the temporal occurrence of lift and first lick relative to time of waterspout contact for the 5 waterspout locations. Median lift-to-contact duration for P1-P5: 179.2, 158.3, 195.8, 320.5, 379.2 msec; data from 1167, 1199, 1136, 1121, 1077 trials of 70 sessions from 25 mice.

Figure 2. Cortical network activity tracks RGD movement progression.

(A) Schematic of wide-field calcium imaging in head-restrained mice reaching with left forelimb. A violet channel was used as the control for the GCaMP channel (blue) to extract calcium activity. sCMOS camera, scientific complementary metal–oxide–semiconductor; LED, light-emitting diode.

(B) Four genetically- and projection-defined projection neuron (PN) types. The intratelencephalic (IT) class includes IT^{Cux1} and IT^{PlxnD1} types; the extratelencephalic (ET) class incudes the pyramidal

tract (PT^{Fezf2}) and corticothalamic (CT^{Tle4}) types. Emx1 marks all PNs. Schematic (right) indicates cortical layers.

(C) Dorsal view of the neocortex showing contralateral regions of interest (ROIs) identified by thresholding of calcium fluorescence change during RGD (see Methods). 1) MOs-c: secondary motor cortex central region; 2) MOp-ul: forelimb primary motor cortex; 3) MO-orf: orofacial motor cortex; 4) SSp-ul: anterior-lateral forelimb somatosensory cortex; 5) Prt: parietal cortex; 6) SSp-bfd/un: anterior part of barrel field and the unassigned region. Dashed lines indicate the boundaries of brain regions registered to Allen CCF. Homotypic ROIs on the ipsilateral hemisphere are not shown. Black squares indicate the center of each region. +, Bregma. Scale, 0.5 mm.

(D) Calcium fluorescence change of different PN populations during water delivery (water), reach, grasp/withdraw and drink; data was pooled from directional reaching for waterspout locations P1-P5. Arrows point to the 6 ROIs. Cortex-wide calcium activity was registered to the Allen CCF and normalized to the max activity change. Average of 9 sessions from 5 mice for PN^{Emx1}; 7 sessions from 4 mice for IT^{Cux1}; 11 sessions from 4 mice for IT^{PlxnD1}; 12 sessions from 6 mice for PT^{Fezf2}; 10 sessions from 5 mice for CT^{Tle4}. The same animals were used in subsequent panels of this figure.

(E) Average cortex-wide calcium activity at the five target locations during the whole RGD process. Note the increase of ipsilateral (left) hemisphere activity as target moved from P1 to P5.

(F) Performance of a generalized linear encoding model (cross-validated variance explained, cvR^2), in which a set of movement variables related to the reaching forelimb was used to predict calcium activity change. Warmer color indicates higher performance in explaining activity with forelimb movement.

(G) Peak activity for all ROIs in ipsilateral and contralateral hemispheres across target locations.

Error bars: SEM.

Figure 3. MOs-c is necessary for the orderly progression and coordination of RGD constituent actions.

(A) ChR2-assisted photoinhibition screening of cortical areas by closed-loop activation of inhibitory interneurons in *Pvalb-ires-Cre;Ai32* mice. Inhibition of the contralateral 1) MOs-c, 2) MOp-ul, and 4) SSp-ul, but not other areas, impaired RGD as reflected by the decrease of supination probability after lift. Colors represent changes in supination probability between inhibition and control trials within 2 seconds after lift; blue indicates occurrence probability decrease. Data from 5 mice. *, significant decrease, Kolmogorov-Smirnov test.

(B) Changes in occurrence probability of RGD constituent movements between inhibition and control trials for each contra- and ipsi-lateral cortical node. Performance of component movements, reach, grasp, and withdraw, was quantified with the probability of successful waterspout aiming, contact, and full hand supination, respectively. The heatmap shows the average difference across 5 mice; blue indicates decrease..

(C) Perturbation of RGD progression in control (left) and inhibition (right) trials upon contralateral MOs-c inhibition. 460/552 lifted control trials and 206/434 lifted inhibition trials finished the RGD action sequence from 5 *Pvalb-ires-Cre;Ai32* mice.

(D) Prolonged inhibition of MOs-c (turquoise) PNs by locally expressing Cre-dependent inhibitory opsin GtACR1 (AAV-DIO-GtACR1) in *Emx1-Cre* mice.

(E) Decrease in the occurrence probability of constituent movements upon prolonged contra-, ipsi-, and bi-lateral inhibition of PN^{Emx1}. Blue indicates decrease. A total of 14 sessions from 7 PN^{Emx1} mice. Data from the same mice was shown in all following panels.

(F) Target location-dependent reduction of aim probability of reaching with prolonged PN inhibition. Aim probability was calculated from all trials with successful lift in each session. ANOVA with inhibition, target location, and subject as factors. Inhibition statistic $F_{1,56} = 19.1$, $p < 0.001$; inhibition × target $F_{4,56} = 5.06$, $p < 0.01$.
(G) Location-dependent reduction of target contact probability of reaching upon prolonged PN inhibition. Target contact probability was from all trials with successful lifts. ANOVA, inhibition $F_{1,56} = 27.89$, $p < 0.001$; inhibition × target $F_{4,56} = 2.99$, $p < 0.05$.
(H) Attenuation of hand opening for grasping with prolonged PN inhibition, shown as average digit opening size at advance endpoint prior to grasping. Wilcoxon rank sum test, $p < 0.01$.
(I) Reduction of supination probability for hand lick after successful grasp with prolonged PN inhibition. ANOVA, inhibition $F_{1,56} = 10.49$, $p < 0.01$; inhibition × target $F_{4,56} = 0.46$, $p > 0.05$.
(J) Schematic of hand-mouth coordination during hand lick for drinking.
(K) Hand-mouth movement profiles of an exemplar trial. Hand upward position, digit open size, mouth open area, and lick onset variables were indicated. Time scale, 100 msec.
(L) Increased premature lick probability during inhibition. ANOVA, inhibition $F_{1,56} = 37.55$, $p < 0.001$.
(M) Hand posture at all lick onsets as represented by hand rotation angle during inhibition. Angle 0 and π indicate palm facing downward and upward, respectively. 13734 control and 9186 inhibition licks. Two-sample Kolmogorov-Smirnov test, $p = 0.0378$.
(N) Increased variance in hand position relative to the mouth upon tongue protrusions during inhibition. ANOVA, inhibition $F_{1,56} = 11.85$, $p < 0.01$.
(O) Decreased coherence between hand upward-downward movement and mouth open-close movement during drinking. ANOVA, inhibition $F_{1,494} = 12.96$, $p < 0.01$.
Error bars/shading: SEM. *, p < 0.05; **, p < 0.01; ***, p < 0.001.

Figure 4. MOs-c neural activity correlates with action relationships and tracks RGD movement progression.
(A) Schematic of electrophysiological recording in cMOs-c (orange dot) during RGD.
(B) Visualization of the 7 identified clusters of neural activity patterns. Each dot represents a neuron. Total 1655 neurons from 111 sessions.
(C) Heatmap representation of the 7 clusters of z-scored neuronal spiking activity of 1655 neurons. Each row represents the activity pattern of a single neuron at five waterspout locations (P1-P5). Neurons were grouped into 7 clusters as activated (cluster 1-5), not modulated (cluster 6), and inhibited (cluster 7). The three vertical dash lines indicate lift onset, advance endpoint, and last hand lick.
(D) Average activity patterns of all clusters. Note that most activity patterns in cluster 1-4 mark the transition or are sustained across action phases.
(E) Population MOs-c activity as decoded by moment-by-moment forelimb reaching kinematics (orange) and its relationship with the waterspout or the mouth (magenta), but was less corelated with movements of other body parts (blue). Gray boxplots indicate cross-validated performance of null model as control of GLM, averaged across 106 sessions.
(F) MOs-c activity decoded target locations across different action phases, starting before hand lift and showing increasing accuracy during subsequent actions. Cross-validated performance of Naïve Bayes classifiers were averaged across 106 sessions. Gray boxplots indicate performance of shuffled model as control.
(G) Visualization of MOs-c population neural dynamics trajectories that evolve along the first three principal components (PC) of all neurons for each waterspout location. Median lift, advance

endpoint, and first hand lick time points are indicated by triangles, circles, and crosses, respectively.

Figure 5. Distinct coding properties and population neural trajectories of MOs-c PN types during RGD

(A) Projection-defined intratelencephalic (IT) and extratelencephalic (PT, CT) neuron types. The same color code is used in subsequent panels.
(B) Activity tuning by target location from example IT^{Cux1}, IT^{PlxnD1}, PT^{Fezf2}, and CT^{Tle4} neurons. Vertical dashed lines: advance endpoint.
(C) Spiking activity of all identified neurons aligned to advance endpoint (vertical dashed line). IT^{Cux1}, IT^{PlxnD1}, PT^{Fezf2}, and CT^{Tle4} neurons are arranged from top to bottom. Each row represents activity across five target locations. For each target location, z-score normalized activity from 1 second before to 1 second after the advance endpoint is shown. Within a neuron type, individual neurons are sorted by their cluster membership (see Figure 4C).
(D) Average spiking activity of all identified neurons aligned to advance endpoint (vertical dashed lines). Note the longer peak latency and more sustained firing in CT^{Tle4} than PT^{Fezf2}. Last row shows merged activity traces of all types.
(E) Proportion of identified PNs that were assigned to each activity cluster identified in Figure 4C. Note the different membership assignment of PT^{Fezf2} and CT^{Tle4} neurons. Circle size indicates the proportion of neurons. n.s., not significantly modulated.
(F) Different PN types show highly distinct population neural trajectories. Whereas PT^{Fezf2} population dynamics display target location modulated trajectories that resemble those of the whole MOs-c population (Figure 4G), CT^{Tle4} trajectories show less target dependency, and IT^{Cux1} and IT^{PlxnD1} show no clear population trajectory. Median lift, advance endpoint, and first hand lick time points are indicated by triangles, circles, and crosses, respectively.

Figure 6. Distinct roles of MOs-c PN types in action sequence and coordination during RGD

(A) Schematic of the prolonged inhibition of MOs-c in four PN types.
(B) Heatmap showing the average change in occurrence probability of RGD constituent actions upon prolonged inhibition of each PN type compared with control trials. Blue indicates decreased occurrence probability. 7 sessions from 6 IT^{Cux1} mice; 11 sessions from 6 IT^{PlxnD1} mice; 14 sessions from 7 PT^{Fezf2} mice; 10 sessions from 7 CT^{Tle4} mice.
(C) Significant decrease of waterspout grasp probability upon PT and CT inhibition. ANOVA. IT^{Cux1} inhibition statistic $F_{1,28} = 0.33$, $p > 0.05$; IT^{PlxnD1} inhibition $F_{1,44} = 0.93$, $p > 0.05$; PT^{Fezf2} inhibition $F_{1,56} = 27.89$, $p < 0.001$; CT^{Tle4} inhibition $F_{1,40} = 8.04$, $p < 0.05$. Note the target location-dependent impairment in PT^{Fezf2} (inhibition × target $F_{4,56} = 2.99$, $p < 0.05$) but not in CT^{Tle4} (inhibition × target $F_{4,40} = 1.24$, $p > 0.05$).
(D) Changes in digit open size at advance endpoint between control and inhibition. Wilcoxon rank sum test.
(E) Decrease of supination probability after grasp during PT and CT inhibition. ANOVA. IT^{Cux1} inhibition $F_{1,28} = 0.76$, $p > 0.05$; IT^{PlxnD1} inhibition $F_{1,44} = 2.49$, $p > 0.05$; PT^{Fezf2} inhibition $F_{1,56} = 12.03$, $p < 0.01$; CT^{Tle4} inhibition $F_{1,40} = 7.66$, $p < 0.01$.
(F) Increased variance of hand position upon tongue protrusion with PN type inhibition. ANOVA. IT^{Cux1} inhibition $F_{1,28} = 2$, $p > 0.05$; IT^{PlxnD1} inhibition $F_{1,44} = 1.47$, $p > 0.05$; PT^{Fezf2} inhibition $F_{1,56} = 7.23$, $p < 0.05$; CT^{Tle4} inhibition $F_{1,40} = 8.25$, $p < 0.05$. Note the target-dependent

impairment in PT^{Fezf2} (inhibition × target $F_{4,56}$ = 4.86, $p < 0.01$) but not in CT^{Tle4} (inhibition × target $F_{4,40}$ = 2.36, $p > 0.05$).

(G) Hand posture at all tongue protrusions as represented by hand rotation angle upon prolonged inhibition of MOs-c PN types. Note the lower hand rotation score upon CT^{Tle4} inhibition. Data from different animals of the same neural type was pooled. Two-sample Kolmogorov-Smirnov test.

Error bars: SEM. *, p < 0.05; **, p < 0.01; ***, p < 0.001; n.s., not significant.

Figure S1. Characterization and quantification of RGD behavior. Related to Figure 1.

A) Exemplar video frames (front view) showing hand locations in a representative reach to P4 at different time points during the reach.
B) Schematic of the measurement of hand rotation direction (black vector) and finger pointing direction (orange vector) as represented by key points on left digits. Hand postures at lift, advance endpoint and supinate for licking are shown.
C) Quantification of waterspout aiming score (cos(δ)).
D) Quantification of digit open size, quantified as the length of hand rotation vector.
E) Schematic for quantifying the hand rotation score (cos(θ)) relative to the horizontal vector (y). Besides wrist rotation movement, the hand rotation score is also a reflection of the palm-facing direction. Value 1 means the palm is facing upward and -1 indicates the palm facing downward. 0 means left hand is facing the right side, which is often the case at the reach endpoint for grasping.
F) Schematic for quantifying hand to mouth distance (d) and hand supination vector (s) upon lick. Hand was maintained close to the mouth and supinated during tongue protrusions.
G) Measurement of mouth opening size (black triangle).
H) Probability distribution of the temporal order of forelimb, hand, and oral movements. Results of 3924 trials from 99 sessions in 33 mice reaching for P2.
I) Exemplar movement variables at waterspout positions P1, P3 and P5.
J) –(O) Target position modulation of lift latency relative to water delivery (J), grasp probability after lift (K), path length (L), target aiming score difference between aim onset and aim end (M), probability of significant acceleration peaks during aim and advance (N), premature lick before waterspout contact (O). 70 sessions from 25 mice.

Figure S2. PN type-based cortex-wide calcium activity during RGD. Related to Figure 2.

A) Neocortical areas superimposed on a brain image registered to Allen CCF. MOB, main olfactory bulb; MOs, secondary motor cortex; MOp, primary motor cortex; RSP, retrosplenial cortex; SSp, primary somatosensory cortex; tr, trunk; ll, lower limb; ul/un, upper limb and unknown region; orf, orofacial; bfd, barrel field; VIS, visual cortex. White dots outline the cortex; yellow horizontal line shows olfactory bulb and neocortex boundary; yellow vertical line shows the midsagittal suture. See Methods.
B) Coronal brain sections showing the laminar pattern of PN types labeled by driver lines crossed to reporter mice, for neuronal types see Figure 2B.
C) Simultaneous immunostaining of Cux1 (red) and Tle4 (blue), with mRNA in situ hybridization of *Fezf2* (magenta) in *PlxnD1;Ai148* (green) mice, showing the spatial distribution patterns of PN types. Scale bar, 100 μm.
D) RGD behavior characterization after training at location P2. Top, lift and first hand-lick probability (Prob.) distribution relative to waterspout contact. 3651 trials in 99 sessions from 33 mice. Bottom, average lick frequency aligned to waterspout contact.
E) Average sequential activity frames with 200 msec interval centered on waterspout contact (black box) during RGD from P2. Average of 9 sessions from 5 mice for PN^{Emx1}; 7 sessions from 4

mice for IT^{Cux1}; 11 sessions from 4 mice for IT^{PlxnD1}; 12 sessions from 6 mice for PT^{Fezf2}; 10 sessions from 5 mice for CT^{Tle4}.

F) Average temporal calcium activity traces aligned to waterspout contact (vertical gray dash line) for different areas and PN populations when RGD from P2. Error shading: SEM.

Figure S3. Target modulation of cortical network activity during RGD. Related to Figure 2.

A) Details of the generalized linear encoding model (GLM). Original calcium dynamics (***Y***) is decomposed into a "spatial" factor ***U*** and a "temporal" factor ***V*** using singular value decomposition (SVD). 10 behavior variables were constructed as ***X*** regressors to predict the temporal components ***V*** with a GLM (***V*** ~ ***BX***). Ridge regularization and cross-validation were used to avoid overfitting. See Methods.

B) Summary and comparison of regions of interest (ROIs) from different neuron types. Top, ROIs by thresholding of normalized GLM performance from Figure 2F. Colors correspond to different PN types. Bottom, ROIs by thresholding of normalized activity from Figure 2E from 5 different waterspout locations.

C) Quantification of the performance of the GLM encoding model for different ROIs. Different colors correspond to different neuron populations.

D) Box plot of target location decoding accuracy of PN type (colored) activities in different ROIs in ipsilateral and contralateral cortex.

E) Comparison of average calcium activity between waterspout location P5 (color traces) and location P1 (gray traces) of ROIs. Arrows indicate increased ipsilateral activities across ROIs and PN types when reaching for P5 over P1. Shaded areas indicate SEM.

Figure S4. Photoinhibition screening across activity nodes and direct inhibition of PNs in MOs-c during RGD. Related to Figure 3.

A) ChR2-assisted closed-loop photoinhibition of cortical areas during RGD. Crosses ('x') represent the center of cortical regions as obtained in previous publications. From anterior to posterior: anterior lateral motor cortex (ALM; Guo, 2014), rostral forelimb area (RFA; Tennant, 2011), rostal forelimb orofacial area (RFO; An, 2023) and caudal forelimb area (CFA; Tennant, 2011) overlaps with primary motor cortex for the upper limb (Sauerbrei, 2020; Muñoz-Castañeda, 2022). Black squares indicate the center of region of interest. +, bregma; scale, 0.5 mm.

B) Distribution of supination latency relative to hand lift in closed-loop inhibition and control trials of each node in contralateral and ipsilateral hemisphere. All trials from 5 *Pvalb-ires-Cre;Ai32* mice were pooled. Two-sample Kolmogorov-Smirnov test.

C) Perturbation of RGD with MOs-c photoinhibition. A control trial is shown along with an inhibition trial as light was triggered during the advance phase. In the inhibition trial, advance was aborted, and the hand returned to the start position (turquoise trajectory), following which another failed attempt occurred. Constituent movements are color coded.

D) Closed-loop inhibition of cMOs-c (contralateral MOs-c) impaired movement progression represented as ethogram from an example session. Actions are color coded. The relative onset and duration of inhibition light for each trial are indicated with light cyan shades. Note inhibition attenuates the completion of the RGD movement.

E) Impairment of movement profiles with closed-loop MOs-c inhibition. Exemplar hand movements in relation to target, hand rotation, and licks from several consecutive control (gray) and inhibition (tortoise) trials are shown. Movement profiles are normalized. Note that upon termination of inhibition, animals immediately resumed and completed the action sequence of RGD.

F) Closed-loop cMOs-c inhibition resulted in increased hand reversals during reaches, decreased target contact after lift, supination after grasp, and hand lick after grasp. Data from 5 mice with 2-sample Kolmogorov-Smirnov test.
G) Viral expression of the inhibitory opsin GtACR1 in MOs-c of an *Emx1-Cre* mouse. Scale, 1 mm.
H) Attenuation of aim and grasp after lift, and supinate and lick after grasp upon closed-loop inhibition of cMOs-c PNs. Trials pooled from 5 PN^{Emx1} mice.
I) Slight decrease of lift probability during reaching for contralateral targets upon prolonged cMOs-c PN inhibition. A total of 14 sessions from 7 PN^{Emx1} mice. Same mice for all following panels. ANOVA, inhibition statistic $F_{1,56} = 10.74$, $p < 0.01$; inhibition × target $F_{4,56} = 4.46$, $p < 0.01$.
J) No significant effects on RGD actions when inhibiting iMOs-c (ipsilateral MOs-c) PN versus control; from 4 PN^{Emx1} mice. Wilcoxon rank sum test for digit opens size; ANOVA for others.

Error bars: SEM. *, p < 0.05; **, p < 0.01; ***, p < 0.001; n.s., not significant.

Figure S5. Dynamic activity tuning to RGD action phase and target location in MOs-c neurons. Related to Figure 4.

A) Raw activity traces of simultaneously recorded neurons using a linear silicon probe from an exemplar trial. Note the different spike patterns at different cortical depths.
B) Examples activity patterns of three simultaneously recorded neurons shown by peri-event histograms and raster plots aligned to the advance endpoint. Trials are sorted by the duration between hand lift (orange dots) and advance endpoint (purple solid line). Black ticks: first hand lick of each trial.
C) Neuronal activity at different neocortical depths relative to hand lift (0) from an example recording session.
D) Fraction of 921 activated neurons that peaked at different action phases during RGD. Note that more neurons showed peak activity during reach, grasp, and withdraw than during drink.
E) Spiking activity aligned to advance endpoint for all recorded neurons sorted by peak activity latency. Each row represents one individual neuron activity across five waterspout locations. A total of 1655 neurons from 111 sessions, of which 921 neurons were activated, 269 neurons were not significantly modulated, and 465 neurons were inhibited. Vertical dash lines represent advance endpoint (time 0).
F) Increase of movement encoding as the depth of recorded neurons increases. $R^2 = 0.125$, $p = 1.05e\text{-}41$. Movement encoding of individual neurons was represented by the deviance explained by a Poisson-GLM. See Methods.
G) Increase of peak activity as the depth of recorded neurons increases. $R^2 = 0.0653$, $p = 3.40e\text{-}15$.
H) Target modulation index negatively correlates with cortical depth. $R^2 = 0.0822$, $p = 6.97e\text{-}19$.
I) Transition probability of neuron activation state across action phases from reach to grasp-withdraw during RGD. Note the higher probability in maintaining either increased or decreased spiking across the two action phases.
J) Computational clustering of all activated neurons using non-negative matrix factorization across a range of cluster numbers. Cluster stability is represented with bootstrapped adjusted rand index (ARI). 5 clusters show the highest cluster stability. See Methods.
K) Sorted membership consensus matrix shows a strong within-cluster consensus and between-cluster antagonism.
L) Tight correlation between exemplar raw behavior profiles (gray) and GLM-predicted traces (orange) using simultaneously recorded spiking activity from an exemplar recording session. Behavior profiles include forward, upward, lateral hand position, and hand position in relation to waterspout or mouth. Spikes were binned in 20 msec bins to predict movement kinematics using GLM.

M) Performance of GLM in predicting hand-target distance is dependent on the number of simultaneously recorded units. y = 0.1861+0.0051x, R^2 = 0.262, p = 3.75e-8.

Figure S6. Optical tagging of PNs. Related to Figure 5.

A) Peri-event activity (top) and raster (bottom) representation of spiking activity of an optically tagged exemplar neuron aligned to blue light pulse onset. Inset shows the average waveform of light-evoked (blue) and regular spontaneous (gray) spikes. Vertical dashed line indicates blue light pulse (blue bar on top) onset. Note there are fast (within 10 msec) and delayed (between 20-40 msec) activations of spiking.

B) Light-evoked spike reliability, jitter, and latency of all neurons (gray and blue circles). Blue circles indicate optogenetically tagged neurons with high reliability, low light evoked spiking latency, and low latency jitter.

C) Light-evoked spiking activity of all identified neurons, including 22 IT^{Cux1}, 17 IT^{PlxnD1}, 63 PT^{Fezf2}, and 30 CT^{Tle4} neurons, arranged from top to bottom. Each row of the heatmap represents the z-score normalized activity of a neuron. Vertical dashed line indicates light pulse (blue bar) onset.

D) Average light-evoked activity of all tagged neurons across different neuron types. Right panel is a zoomed-in view of the gray area in the left panel, which indicates the delayed modulation upon brief light pulses. Horizontal dashed lines: 0 z-score.

E) Boxplots of recording depth, jitter and latency of light-evoked spike, and baseline firing rate of different PN types. Median ± interquartile range. Horizontal lines in boxplots indicate 75%, 50%, and 25% percentile. Whiskers represent data point span to 90% or 10% percentile.

Figure S7. PN-type specific behavioral correlates and properties of population dynamics. Related to Figure 5.

A) Hand movement trajectory with neural activity (colormap) superimposed from exemplar neurons of four different PN types. X, forward; Y, lateral; Z, upward positions. Scale, 5 mm.

B) Peri-event histogram and raster plot from optically tagged IT^{Cux1}, IT^{PlxnD1}, PT^{Fezf2} and CT^{Tle4} exemplar neurons. Spikes are aligned to advance endpoint (purple vertical line). Trials are sorted by the onset of hand lift (orange ticks). Vertical dashed line: advance endpoint.

C) Average firing rate of the four PN types at different RGD action phases.

D) Distribution of movement encoding performance of individual neurons. Movement encoding performance of individual neurons was represented by the deviance explained by a Poisson-GLM. PT^{Fezf2} and CT^{Tle4} comprise a higher fraction of neurons that reliably and robustly encode movement.

E) Target modulation index of distinct PN types across different action phases.

F) Cumulative proportion of variance explained after PCA decomposition of different PNs. Bootstrapped component number to explain 80% of the variance of 16 neurons for each neuron type: 4 for IT^{Cux1}, 3 for IT^{PlxnD1}, 3 for PT^{Fezf2} and 2 for CT^{Tle4} as indicated by the dashed line.

G) Smoothness of the population neural trajectory. Note the higher smoothness (lower jerkiness) in PT^{Fezf2} and CT^{Tle4} populations. Equal numbers of neurons (16 from each neuron type) were randomly selected for the analysis to eliminate the impact of neural number. The following metrics of population neural trajectory were all bootstrapped for comparison among neuron types.

H) Response distance between pairs of neural trajectories at advance endpoint for different PN types. Heatmap represents bootstrapped average Euclidean distance. Note overall larger response distance in IT^{Cux1} and the overall smaller response distance in CT^{Tle4}. Also note lower response distance between closer waterspout locations.

I) Dissimilarity in geometric shape of different pairs of population neural trajectories of different neuron types. Note overall lower shape similarity (higher dissimilarity) in IT^{Cux1} and the overall higher shape similarity (lower dissimilarity) in CT^{Tle4}.

Figure S8. MOs-c PN-type specific contribution to RGD. Related to Figure 6.

A) Cortical, striatal, and thalamic projections from different MOs-c PN types. IT neurons show projection to contralateral cortex and no projection to thalamus. Note the rare projection of IT^{Cux1} to striatum as compared with that of IT^{PlxnD1}. PT^{Fezf2} and CT^{Tle4} both project to thalamus, with CT^{Tle4} projections restricted only to thalamus. Scale, 1 mm.

B) Heatmap summary of the change of RGD constituent action occurrence-probability upon closed-loop inhibition of different PN types in MOs-c.

C) Reduction of occurrence-probability of contact for grasp after lift (left), supinate after grasp (middle), and lick after grasp (right) upon closed-loop inhibition of MOs-c IT^{PlxnD1}, PT^{Fezf2} and CT^{Tle4} neurons. Two-sample Kolmogorov-Smirnov test.

D) Impairment of waterspout grasps after lift in prolonged inhibition trials of different PN types. Trials are pooled from 6 IT^{Cux1} mice; 6 IT^{PlxnD1} mice; 7 PT^{Fezf2} mice; 7 CT^{Tle4} mice.

E) Prolonged MOs-c inhibition in IT^{PlxnD1}, PT^{Fezf2}, and CT^{Tle4} increased the probability of premature lick. Wilcoxon rank sum test.

METHODS

Mice

Animal care, use, surgical and behavioral procedures conformed to the guidelines of the National Institutes of Health. The experiments were approved by the Institutional Animal Care and Use Committee of Cold Spring Harbor Laboratory and Duke University. Experiments were conducted with 8-week- to 16-week-old male and female mice. The number of animals used in each experiment is noted in the corresponding section. Mouse strains were: *Emx1-Cre* (JAX#005628), *Pvalb-IRES-Cre* (JAX#017320), *Tle4-CreER* (JAX#036298), *Fezf2-CreER* (JAX#036296), *Cux1-CreER* (JAX#036300), *PlxnD1-CreER* (JAX#036294), *Ai148D* (JAX#030328), *Ai32* (JAX#024109). All mouse driver lines were bred with reporter strains for calcium imaging or electrophysiological recording except for *Emx1-Cre* mice. They were injected retro-orbitally with AAV-PHP.eB-CAG-DIO-GCaMP7f (2.46×10^{13} GC/mL, Addgene) at postnatal days 14 for as they failed to breed with *Ai148D* mice. Mice were housed in groups of up to five mice per cage, in a room with a regular 12/12 light/dark cycle. After surgery, mice were housed in a new home cage individually or with familiar groups for at least one week prior to experiments.

Surgery

Materials, including instruments used in surgery, were sterilized, and stereotaxic surgery was performed using aseptic techniques. Surgical anesthesia was maintained using 1%-2% isoflurane via inhalation. The analgesic drug ketoprofen (5 mg/kg, subcutaneous), with an effect lasting up to 24 hours, was administered prior to the beginning of the surgery. The local anesthetic lidocaine was administered subcutaneously at the intended incision site (2-4 mg/kg). Body temperature was maintained at 37° C using a feedback-controlled heating blanket. Eye ointment was applied to prevent the eyes from drying.

After disinfecting with betadine solution (5-10%) and ethanol (70%), a small incision of the scalp was created to expose the skull. A titanium flat head post was implanted for head restraint experiments. For wide field calcium imaging, the skull was cleaned thoroughly with saline, and a thin layer of cyanoacrylate glue (Zap-A-Gap CA+, Pacer Technology) was applied on the skull to clear the bone. After the cyanoacrylate glue was cured, cortical blood vessels were clearly visible. Then, a circular flat head post was attached to the skull using dental cement (C&B Metabond, Parkell; Ortho-Jet, Lang Dental) while leaving most of the dorsal cortex exposed. For inhibition screening experiments, a thin skull preparation [99] was used in *Pvalb-IRES-Cre*;*Ai32* mice. Clear low toxicity silicone adhesive (KWIK-SIL, World Precision Instruments) was applied on the dorsal cortex as protection to dust and scratches. Animals were allowed to recover for one week after surgery before the experiments.

Viral vector injections were performed using a microsyringe pump (Nanoliter 2010 Injector, WPI). A SMARTouch controller (WPI) delivered the virus at a rate of 46 nL/min. For optogenetic manipulation, a 300-500 nL volume of virus (AAV_{DJ}-CAG-DIO-GtACR1-EYFP, 2.43×10^{13}

GC/mL, Vigene Biosciences) was injected into the cortex at the central premotor cortex (MOs-c) region. Injection coordinates (AP/ML) of MOs-c were +1.6/±1.4 mm. After injection, the glass pipette was left in place for ten minutes and then slowly withdrawn at a speed of 50 μm per min. Optical fibers (200 μm, 0.37 NA; RWD Life Science Inc.) fitted into an LC-sized ceramic fiber ferrule were implanted. After 12 to 21 days of waiting time for post-surgery recovery and viral expression, the animals were used in experiments.

Tamoxifen induction

Intraperitoneal injection of tamoxifen (two 100 mg/kg injection at 20 mg/ml, prepared in corn oil) was performed to induce Cre recombinase expression for reporter crossed *CreER* mice. The first induction was on the day of weaning and the second induction was one week later. For Cre recombinase induction in virus injected *CreER* mice, two injections of tamoxifen were administered intraperitoneally on day 1 and day 3 after virus injection (day 0).

Head-restrained reach-to-drink

The day before behavioral training, animals were weighed and moved to a new cage with new bedding and food while restricting access to water. Animals received supplemental water to meet a daily water need of 1 mL to maintain body weight >80% of the initial weight, monitored by daily weighing and evaluation.

The reach for water task [45] was controlled in real-time with MATLAB (MathWorks). The data acquisition board (USB-6351; National Instruments) communicates between the software and hardware (piezo sensors, water valves, and linear actuators). Two high-speed USB cameras (FL3-U3-13S2C-CS or BFS-U3-04S2C-CS; FLIR) acquired video data from the front and left side of the mouse. The cameras were synchronized and calibrated to enable three-dimensional infrared recording of the animal's forelimb and orofacial movements. Simultaneous acquisition and storage of the video at 240 frames per second (fps) at a resolution of 640×480 pixels was achieved using Bonsai such that trial information and touch sensor data were shared in real-time with MATLAB and the cameras.

Mice were trained to reach for water in two phases, within which they were required to reach a hit rate > 80%. In phase 1, the waterspout was fixed on the left side of the animal's snout, and in phase 2, it was moved to one of 5 equidistant locations (identified as left P1, P2, center P3, right P4, P5) with P3 centered and each location approximately 3 mm apart in front of the animal's nose. Phase 1 training consisted of 1 session each day for 3 days, with 100 trials per session. Pretraining to reach involved placing a waterspout (made from an 21-gauge needle and providing a drop of sucrose solution, 10% w/v, 20~50 μL) ~3 mm to the left side of the snout midline. The waterspout tip was horizontally aligned with the upper point of animal's mouth, which is about 4 mm below the tip of nose. Animals were trained to use left hand to grasp the water drop with right limb blocked. The waterspout tip was initially close to the mouth but then gradually moved away to 3-5

mm from animal's mouth. Water was delivered at a random inter-trial interval (12-20 s). The random duration was long enough for the animal to replace its hand to the starting position after consumption. The hand was 20-30 mm posterior and downward from the waterspout tip at the starting position. A piezo sensor detected waterspout contact events. If the animal failed to reach within 8 seconds after water delivery, a new trial began after a 30 second timeout. For phase 2 training, a linear actuator (L16-R Miniature Linear Servo for RC; Actuonix Motion Devices) would move the waterspout to one of the 5 locations.

Wide-field calcium imaging

Wide-field calcium imaging [41,60,100], was performed with an inverted tandem-lens macroscope in combination with a scientific complementary metal-oxide semiconductor (sCMOS) camera (Edge 5.5, PCO). The top lens had a focal length of 105 mm (DC-Nikkor, Nikon) and the bottom lens 85 mm (85M-S, Rokinon), resulting in a magnification of ×1.24. The total field of view was 12.4 mm by 10.5 mm with a spatial resolution of ~20 μm/pixel. To capture GCaMP fluorescence, a 525 nm band-pass filter (#86–963, Edmund optics) was placed in front of the camera. Using alternating excitation light at two different wavelengths, calcium-dependent fluorescence was isolated and corrected for intrinsic signals (for example, hemodynamic responses). Excitation light was projected onto the cortical surface using a 495 nm long-pass dichroic mirror (T495lpxr, Chroma) placed between the two macro lenses. The excitation light was generated by a collimated blue LED (470 nm, M470L3, Thorlabs) and a collimated violet LED (405 nm, M405L3, Thorlabs) that were coupled into the same excitation path using a dichroic mirror (#87–063, Edmund optics). The alternating illumination between the two LEDs and the acquisition by the imaging camera were controlled by an Arduino Uno R3. The camera ran at 50 fps, producing one set of frames with blue excitation and another set with violet excitation, each at 25 fps. The exposure state of each frame was recorded. Excitation of GCaMP at 405 nm resulted in non-calcium-dependent fluorescence, allowing isolation of the true calcium-dependent signal by subtracting fluorescence changes in violet frames from the blue illumination frames by regression, as detailed below. Subsequent analyses were based on this differential signal at 25 fps. A common synchronization signal is shared between behavioral cameras and imaging camera to align real-time behavior and neural data.

Optogenetic manipulation

The open-source guide was used to achieve real-time and closed-loop control based on markerless hand position tracking [101]. Briefly, real-time reach behavior was monitored using a USB camera (Flea3; Point Grey) on the left side of the mouse, ipsilateral to the reaching forelimb. A trained deep neural network with ResNet-50 model (the same one used for behavior analysis) for the side view video was embedded in a custom Bonsai workflow to trigger optogenetic stimulation based on real-time detection of hand position (https://github.com/bonsai-rx/deeplabcut). Low-latency control of light was achieved with videos capturing at 25 fps and a resolution of 640 × 480 pixels

on a Windows workstation equipped with a GeForce RTX 2080 Ti GPU (NVIDIA). A 200 μm optical fiber delivered 473 nm blue (SSL-473-0100-10TM-D, Sanctity Laser) or 532 nm green (SSL-532-0200-10TM-D, Sanctity Laser) light.

Inhibition of neural activity of different cortical areas was achieved with *Pvalb-IRES-Cre*;*Ai32* mice that allow optogenetic activation of local PV interneurons for an inhibition screening [65]. Blue light pulses (5 msec, 50 Hz, 473 nm) were triggered in 50% of reach trials, as the animal's real-time hand position crossed a predefined threshold in a closed-loop manner. The light spot size was restricted with a 200 μm optical fiber, with its tip directly contacting the thinned skull. The fiber ferule was positioned with an MP-285 micromanipulator (Sutter Instrument). Light intensity at the tip was adjusted to 5 mW. The triggered light was automatically turned off 4 seconds after the water delivery.

For other inhibition experiments, the inhibitory opsin GtACR1 was locally expressed by viral injection and stimulated by green light (532 nm) adjusted to 5-10 mW. Two optogenetic inhibition strategies were applied. 1) Closed-loop reach photoinhibition: the light was on when the hand moved across a predefined position during reaching and off when the hand repositioned at the resting bar. 2) Prolonged inhibition: light was turned on 1 second before water delivery and lasted for the entire trial. Experiments involved either bilateral or unilateral inhibition.

Multielectrode array recording

The surgery procedure was as described in previous sections. To provide a ground reference, an M1 screw connected to a silver wire (A-M systems) was implanted into the skull above the left visual cortex. Before the first recording session, a craniotomy was performed under isoflurane anesthesia. A linear silicon probe was slowly lowered into the cortex with an MP-285 micromanipulator (Sutter Instrument). A clear silicone elastomer (Kwik-Sil, World Precision Instruments) was applied over the craniotomy after the electrode was positioned to the desired position to stabilize the exposed brain. The brain was allowed to settle for 15-30 minutes before recordings began. At the end of the recording session, the probe was retracted and the craniotomy was sealed with silicone elastomer to allow a subsequent session on the following day.

Behavioral videos, light pulses, and electrophysiological data were synchronized with a common synchronization signal. Extracellular spikes were recorded using multielectrode arrays (ASSY-37 H4, Cambridge NeuroTech, or A1×32-5mm-25-177, A4×8-5mm-100-200-177, NeuroNexus). Voltage signals were continuously recorded at 32 kHz from 32 channels of the silicon probe by a Digital Lynx 4SX recording system (Neuralynx). Raw data was collected and saved using Cheetah software. Neuronal activity was band-pass filtered (300-6000 Hz) for real-time visualization. During the first and last five minutes of the recording session, when animals were not engaged in the behavioral task, 473 nm blue light pulses (2 msec or 5 msec duration) at different frequencies (0.1 or 10 Hz) were delivered through an optical fiber over the craniotomy for optical tagging. Light intensity was adjusted based on the amplitude of light evoked response during the process.

RNA *in situ* and immunohistochemistry

After the experiments, animals were euthanized with isoflurane. Using a peristaltic pump they were perfused for 15 min with chilled saline (0.9% NaCl) followed by fixative (4% PFA in 0.1 M PBS) . Brains of *PlxnD1-CreER;Ai148* mice were dissected, fixed, and cut into 10 µm sections. For RNA *in situ* hybridization chain reaction, brain sections were hybridized with Fezf2 probe (Molecular Instruments) in probe hybridization buffer at 37 ^{o}C for 24 hours, washed with probe wash buffer, and incubated with an amplification buffer at 25 ^{o}C for 24 hours in a 24-well plate. For immunohistochemistry after RNA *in situ*, the same brain sections were stained with rabbit anti-Cux1 (1:500, 11733-1-AP, Proteintech) or mouse anti-Tle4 antibody (1:500, Cat#sc-365406, Santa Cruz Biotechnology). Briefly, brain sections were first pretreated with 10% Blocking One (Cat#03953-95, Nacalai Tesque) in PBS with 0.3% Triton X-100 (Blocking solution) at room temperature for 1 hour, then incubated with primary antibody in Blocking solution at 4 ^{o}C overnight. The brain sections were washed with PBS the next day and incubated with Cy3 conjugated donkey anti-rabbit (1:500, Cat#711-165-152, Jackson Immunoresearch) or donkey anti-mouse (1:500, Cat#715-165-150, Jackson Immunoresearch) secondary antibody for 2 hours at room temperature. Brain slices were then mounted on slides for confocal imaging (ZEISS Axio Observer).

Data processing and analysis

Data processing and analyses were performed with MATLAB (Mathworks) or Python, unless otherwise specified. Sufficient reach-to-grasp trials were collected for each condition such that all results could be reproduced robustly. No statistical methods were used to predetermine sample size.

High-speed behavior tracking

The reaching behavior was monitored by two high-speed cameras (Flea 3, Point grey) at 240 fps from both the front and the left side of the animal. The geometric relationship between the two cameras was calculated using the Camera Calibrator app in MATLAB (2020a). Real-time videos of the behavior session were acquired and saved using a customized code in Bonsai (version 2.6) for offline analysis. Two deep neural networks were trained separately using DeepLabCut (2.0) to track the body part positions in front and side views. Neural network training was performed using over 2000 frames (1040 front and 1076 side frames) from 20 different recording sessions of 20 individual animals. Images of different animal colors, body sizes, head-restrained setups, illumination conditions from different behavior phases were included to train a relatively robust network. In total, 18 keypoints in the front view and 22 keypoints in the side view were labeled on each frame. These included the digits, nose, mouth, tongue, waterspout, and the water drop. For the front view, 1030000 training iterations were achieved with a training error of 2.1 pixels and test error of 6.6 pixels. At a statistical confidence level of >0.95, the training error was 2.07 pixels and test error 3.89 pixels. For the side view, 1030000 training iterations produced a training error of 1.74 pixels and test error of 5.04 pixels. With a confidence level cutoff of >0.95, the training error was 1.74 pixels, and the test error was 4.16 pixels. The 3D positions of the hand and waterspout were reconstructed through stereo triangulation. Only samples with a network predicting confidence level of >0.95 were used for analyses. In cases of missing samples, the corresponding samples from a cubic spline were used to fill the trajectory (gaps >100 msec were not used).

Analysis of constituent movement components and action segments of RGD

The RGD behavior is a closed loop act in which mice monitor the target location online using olfactory and tactile cues to direct their reach [45,102]. The movement of reach, grasp and withdraw and their constituent actions were identified using a previously described movement classification scheme [103]. The reach consists of lift, aim and advance segments and directs the hand to the waterspout. The grasp consists of opening and extending the digits and then closing them to purchase a water drop. The withdraw is an egocentric movement in which the hand is supinated after grasp as hand retracts to the mouth for drinking. The following constituent action segments and critical events (e.g. start and end points of an action) of the RGD are featured in the automated kinematic analysis pipeline for pose estimation after keypoint extraction by DeepLabCut.

Lift. The hand was raised from the resting position and partially pronated with the digits collected (i.e. lightly closed and flexed). The 3D reconstructed position of digit 3 was used to represent and track hand movement. The first frame in which the vertical hand speed increased above 75 mm/s (upward) defined lift initiation. Left-hand speed was the absolute value of the derivative of 3D left-hand position. The lift phase consisted of the time series from lift initiation to peak speed.

Aim. The hand was positioned by an elbow-in movement of the upper arm, the digit orientation was rotated and palm aimed toward the waterspout. Hand/digit rotation was characterized using the supination vector that connects the midpoint and tip of digit 3 and digit 4 in the front view (Figure S1B). The vector connecting moment-by-moment hand position and the waterspout position was defined as the range vector (vector ***r*** in Figure S1C). The moment-by-moment direction and amplitude change of range vector reflected the hand movement in relation to the waterspout. The waterspout aiming angular deviation (angle δ in Figure S1C) was the angle between the current finger pointing direction (vector ***p*** in Figure S1C) and the instantaneous range vector direction (vector ***r*** in Figure S1C) at a given moment. A waterspout aiming score was defined as the cosine value of the angular deviation ($\cos(\delta)$ in Figure S1C). Aim completion was defined as the time point when waterspout aiming score is higher than 0.866 ($\delta < \pi/6$). The aiming phase was defined as the time points from first hand peak speed point to aim completion.

Advance. The hand was advanced toward the waterspout by upper arm movement and opening of the elbow with concurrent opening and extension of digits. The advance phase was defined as the time from aim completion to advance endpoint which terminates just before grasp. If an animal failed to reach the waterspout or to grasp water, the hand would usually return to the aim position with the digits closed and flexed for another advance.

Advance endpoint. The vector direction from the initial hand position to the waterspout was the reference direction (dashed line in Figure 1C). The median hand position during the 2 seconds before water delivery (pre-lift position) was used as the initial hand position. Real-time hand to spout distance was calculated as the projection of the range vector onto this reference direction. The furthest reaching point along reference direction defined the reaching endpoint. At this point the hand was positioned adjacent to the waterspout with the digits extended and the palm in a near vertical orientation in preparation for grasping. The length of the rotation vector (Figure S1D) reflects how much the digits were abducted and is defined as digit-open size. The amplitude and direction of the range vector at the reaching endpoint were characterized to represent reaching endpoint accuracy.

Grasp. The hand contacted the waterspout and the digits were closed and flexed to purchase a water drop. Grasp completion was defined as the time point when tips of all 4 digit become invisible (confidence level < 0.5) from the lateral view following hand-open and digit-extended state. The waterspout contact event was identified by the piezo sensor attached to the waterspout.

Withdraw and Supination. The hand was withdrawn to mouth by upper arm movement that lowered the hand to the level of the mouth with the palm in a vertical orientation. It was further supinated by movement at the wrist such that the palm faces upward. Supination was measured as

wrist rotation to a position that the palm faces up with the hand rotation score higher than 0.5 ($\theta < \pi/3$). The rotation score was the cosine value of the direction of the line connecting midpoints of digit 4 and digit 3 in the front view ($\cos(\theta)$ in Figure S1E).

Licking. Tongue protrusion events were identified by the trained deep neural network for front view videos with confidence level > 0.95. Most tongue protrusions occurred after animals successfully grasped the waterspout. Withdraw-to-drink phase was the time points between the advance endpoint and the first hand-lick.

Hand-mouth coordination during drinking. Animals consumed water by licking their left hands. The hand was positioned near the mouth and makes repositioning movements that included digit opening and extending in coordination with licking. The median position of the left and right mouth corners in that session was the mouth position. The distance between moment-by-moment hand position and the reference mouth position in the front view was used as hand to mouth distance. The distance between the hand and tongue (***d*** in Figure S1F) and the hand rotation score (***s*** in Figure S1F) upon tongue protrusion quantified the coordination between hand and tongue movement for drinking.

Other orofacial movements. Prior to reaching, animals detected the water by sniffing and orienting their nose toward the corresponding waterspout location. Moment-to-moment nose displacement movements were obtained by comparing the median value of nose positions with the left side displacement being positive and right displacement being negative. Mouth open and close movements were identified by calculating the area covered by two mouth corners and upper lip in the front view video (Figure S1G).

Replace. After drinking, the hand reversed movement direction, lowered from reaching or from the mouth, and returned to the approximate starting position.

Cortex-wide neural dynamics

Analysis of wide field data used methods described previously [41,100,104]. The landmarks of the dorsal cortex were marked, and the mask was set in the scope of the dorsal cortex from an example frame. Next, control and GCaMP frames were split from raw videos. Each cropped video frame (size 440 × 440) was transformed to a flat array. Images from different trials were then concatenated resulting in a two-dimensional matrix (size $n \times t$, $n = 440 \times 440$, t is time). Denoising was performed with SVD-based method (singular value decomposition). SVD returned 'spatial components' U (of size pixels by components), 'temporal components' V^T (of size components by frames) and singular values S (of size components by components). To reduce computational costs, all subsequent analyses were performed on the product SV^T represented as V_c. Results of the analyses on SV^T were later multiplied with U to recover results back to the original pixel space. The denoising step outputs a low-rank decomposition of $Y_{raw} = UV_c + E$ represented as an $n \times t$ matrix; here UV_c is a low-rank representation of the signal in Y_{raw}, and residual E is considered noise. The output matrices U and V_c are much smaller than the raw data Y_{raw}, leading to

compression rates above 95%, with minimal loss of the visible signal. Finally, an established regression-based correction method isolated a purely calcium dependent signal by subtracting the control channel signal Y_v (405 nm illumination) from the GCaMP channel signal Y_g (473nm illumination). With the corrected values for each trial, the signal values in a time window one second before water delivery were used as the baseline to calculate z-score. All wide field imaging data was registered to the Allen reference mouse brain Common Coordinate Framework (CCF3) using five anatomical landmarks: the left, center and right points where anterior cortex meets the olfactory bulbs, the medial point at the base of retrosplenial cortex, and Bregma, labeled manually for each imaging session during mask setting (Figure S2A).

Encoding model for cortex-wide neural dynamics

Generalized linear encoding models (GLM) with ridge regularization were built to predict the cortex-wide neural dynamics with animal behavior (behavior matrix *M*, size frames by 13) [41]. Cortex-wide neural dynamics were represented by all 'temporal components' V_c (size components by frames) saved after SVD decomposition. The model was fitted using ridge regression with 10-fold cross-validation to avoid overfitting. The regularization penalty was estimated separately for each component of V_c data on the first fold of validation and used the same value for other folds of validation. A new modeled variable V_m of the same size as V_c was predicted using this GLM process and used to compute $Y_m = UV_m$ as predicted pixel-wise neural dynamics. The predicted Y_m was compared with Y_{raw} and pixel-wise explained variance (R^2) was obtained to quantify the cross-validated GLM performance (represented with cvR^2).

To construct the behavioral matrix *M*, movement profiles related to hand, digit, wrist, and orofacial movements were derived from the frame-by-frame labeled points in front and side videos. In addition to the isolated left-hand movements, the hand relationships to the waterspout and to the mouth were included to fully capture the events and their spatiotemporal relationships that constitute RGD. Thirteen analog behavioral variables after kinematic analysis were selected to describe hand, digit, wrist, and orofacial movement during RGD. Those ethologically meaningful behavioral variables were not necessarily orthogonal to each other. Normalized behavior variables were downsampled to match cortex wide activity, combined to make matrix *M* (size frames by 13), and used to predict cortex-wide neural dynamics. Thirteen continuous variables are listed as follows (Figure S1I): 1) forward position of left hand, 2) upward position of left hand, 3) lateral position of left hand, 4) path length moved from left hand onset, 5) moving speed of left hand, 6) digit open size of left hand, 7) moving speed of right hand, 8) nose displacement, 9) mouth open size, 10) supination score of left hand, 11) left hand to mouth distance, 12) waterspout aiming score by digits of left hand, and 13) left hand to waterspout distance. The first 6 variables represent the kinematics of left hand movement. Variables from 7 to 9 depict the movement of other body parts. The last 4 variables (from 10 to 13) reflect the relationship between left reaching hand and target or between left reaching hand and mouth. Reaching forelimb-related variables (1-6 and 10-13) were used to predict calcium dynamics. We also tried predicting neural activity with all

thirteen movement variables. Similar results were observed. No time-shifted versions of movement profiles were tested.

Decoding waterspout locations with calcium dynamics

We trained naive Bayes classifiers to predict waterspout locations using calcium activity from individual trials in each area within a session. Cortex-wide neural dynamics of each session were represented with all 'temporal components' *Vc*. The values from 2 seconds before to 4 seconds after waterspout contact were averaged to obtain a single value for each temporal component representing the activity for that trial. For a session of *k* trials, the predictors would be size *k* by the number of components. The response variable would be *k* rows with class information.

The naive Bayes classifiers were trained using MATLAB (*fitcnb*). Classifiers provided a probabilistic assignment of the input to one of five possible waterspout locations or three classes relative to body sides. The class with the maximum posterior probability was chosen as the predicted class. A cross-validation procedure in which two random trials for each waterspout location were left out were used to train the classifier on the remaining data to predict five waterspout locations. The classifier was then tested on the left-out samples. This procedure was repeated 100 times and the average test performance represented classification accuracy. As a control, the input of validation sample was randomly shuffled to get a shuffled measure. The same prior probability for five waterspout locations was used. Classifiers for body side-based waterspout locations were trained in the same way with the five classes replaced by three waterspout locations relative to body side with different priors.

Cortical network nodes

Masks for regions of interest (ROIs, Figure S3B) were derived by thresholding averaged calcium activity amplitude (Figure 2E) and the GLM encoding model performance (cvR^2) (Figure 2F). All neuron types were considered. Activity amplitude or GLM performance larger than 80% of the maximum value of all pixels on the dorsal cortex was used as the high threshold. Activity amplitude lower than -0.4 z-score was used as the low threshold. ROI 1 (MOs-c, central region of secondary motor cortex, centered at AP/ML: +1.60/±1.37 mm) and ROI 5 (Prt, parietal cortex including part of SSp-ll and SSp-tr, -1.18/±1.63 mm) were overlapping regions with activity amplitude larger than the 80% of maximum activity from Emx1, Fezf2 and Tle4 populations (Figure S3B). ROI 4 (SSp-ul, anterior-lateral forelimb somatosensory cortex and the unassigned region, +0.23/±2.62 mm) was observed in Tle4 activity larger than high threshold and further isolated with GLM performance of Emx1, Fezf2 and Tle4 populations. ROI 3 (MO-orf, orofacial motor cortex in the lateral part of the anterior cortex, +1.67/±2.08 mm) covered the pixels that were lower than -0.4 z-score in Cux1 activity. ROI 6 (SSp-bfd, anterior part of barrel field and nose primary somatosensory cortex, -0.71/±2.69 mm) switched hemisphere as waterspout was moved from ipsilateral to contralateral side of the animal from both Cux1 and PlxnD1 populations

(Figure S3B). ROI 2 (MOp-ul, primary motor cortex for upper limb, +0.41/±1.69 mm) was revealed in the GLM performance of Cux1 population and further verified by the extended Fezf2 activity into MOp on the right hemisphere as compared to the left hemisphere. Because the border of some ROIs overlapped, the coordinates of the center of each ROI relative to Bregma were shown above and used to guide further experiments. The median value of pixels within a 100 μm diameter from the center of a ROI was used to represent its activity.

Waterspout modulation index

Waterspout modulation index reflects the difference in activity amplitude when animals reach for different waterspout locations. Modulation index (*MI*) was calculated as follows: $MI = (A_{max} - A_{min}) / A_{max}$; where A_{max} (maximum activity) and A_{min} (minimum activity) were computed for different waterspout locations. For wide-field imaging data, pixel-wise *MI* values were provided at specific behavioral time points. For electrophysiological data, *MI* values of individual neurons were shown across different behavioral phases.

Spike sorting and optical tagging

Saved raw electrophysiological data was rearranged by the channel depth, median-subtracted across channels and time, and the results were saved in 16-bit binary files for further spike detection and sorting using Kilosort2 software (https://github.com/cortex-lab/KiloSort). A proper probe configuration was created using default parameters for spike detection and sorting. Spikes were further visualized and manually curated in phy2 (https://github.com/cortex-lab/phy). Sorted data was analyzed using custom MATLAB scripts. Several parameters were taken into consideration for cluster quality control: spike shape, average spike firing frequency (> 0.05 Hz), amplitude (> 60 mV), contamination rate (< 0.2), and isolation distance (> 18). Neurons were further curated by only keeping those with correlation coefficient of spike waveform >= 0.85 between random selected spontaneous spikes and light stimulation evoked spikes.

Optical tagging was adopted to confirm the cell type of identified good quality clusters [69,70]. Light pulses (5 msec, 0.1 or 10 Hz) were passed through the optical fiber to stimulate the neurons. The light intensity was lowered to reduce spike jitter. The latency and waveform of first spikes after light pulses within 8 msec were used for the validation of optical tagging. We then computed the reliability of light-evoked spiking within 8 msec from light onset (R), the median latency of triggered spikes after light onset (L), the variance of the latency across different light pulses (J), and a statistical p-value to determine whether the spikes were truly evoked by light stimulation. The p-value was determined by comparing the distribution latencies of light evoked spikes and a bootstrapped distribution of latencies of spontaneous spikes. Only neurons with $R >= 0.3$, $J < 2$ msec, and $p < 0.001$ were considered as optically tagged neurons.

PETH

To compute peri-event histograms (PETH) aligned to different behavioral events, the spikes of individual neurons were binned in 10 msec windows and smoothed with a Gaussian kernel (20 msec standard deviation). Average spike rates across trials were used to represent the activity of individual neurons. All spike rates were z-score normalized to the mean and standard deviation of a pooled distribution of binned baseline activity (1 second before water delivery) across all trials. Each unit had twenty PETHs related to four behavioral events: hand lift onset, advance endpoint, first lick, and last lick, each with five different target locations.

Neurons with significant activity increases or decreases were compared to baseline ($p < 0.05$) during 600 msec around the advance endpoint (300 msec before and 300 msec after advance endpoint) were considered as activated or inhibited neurons, respectively. All other neurons were considered as not significantly modulated. Non-parametric Wilcoxon signed-rank tests were used to compare all trials.

NNMF clustering

Non-negative matrix factorization (NNMF) was used to identify optimal cluster numbers and clustering of temporal neuron activity patterns [14]. To construct the input matrix to NNMF, spikes associated with each neuron were binned with a 20 msec window in each trial and were concatenated across time. NNMF was performed using the MATLAB function *nnmf* using default options. For any rank *k*, the NNMF algorithm groups the samples into clusters. The key issue is to determine whether a given rank *k* decomposes the samples into “meaningful” clusters. For this purpose, an approach to model selection that exploits the stochastic nature of the NNMF algorithm was developed. The NNMF algorithm may or may not converge to the same solution on each run, depending on the random initial conditions. If the clustering of *k* classes is strong, the expectation is that sample assignment to clusters would vary little from run to run. To find the best number of clusters, a range of numbers with bootstrap cross-validation was tested to see what cluster number produced the most consistent cluster membership. In each bootstrap iteration, NNMF with a given cluster number was applied using 90% of the randomly sampled neurons. The extracted activity patterns were used to compute cluster memberships for the other 10% of neurons that were held out. This process was repeated 500 times. The final cluster membership of a neuron was the one that had the highest likelihood of containing that neuron. We ran this method on all activated neurons with the number of clusters set to each value from 4 to 16, and found that 5 clusters achieved the best consistency, quantified as the mean likelihood that a neuron was grouped in the same cluster across all bootstrap iterations.

Encoding and decoding GLM with spikes

To reveal the information encoded by individual neurons, we used Poisson GLM (*glmfit*) to predict the spike count within 20 msec bins with movement profiles (behavior matrix *M*) [105].

For decoding, the spiking activity of simultaneously recorded neurons from each session was used to predict thirteen behavioral variables contained in the behavioral matrix *M* individually with Gaussian GLM (*glmfit*) within a session. Spiking activity was binned and smoothed with a 20 msec window for each neuron. Behavioral profiles were down sampled to match spiking activity. No lagged versions of movement profiles were tested. Either normalized activity or the principal components explaining over 85% of the variance after PCA in each session was used as regressors, which resulted in the same conclusion.

To avoid overfitting, similar procedures were used to perform cross-validation for both encoding and decoding models, in which 25% of samples were left out in a random trial-based manner with trained GLM on the remaining data. The trained model was used on the left-out sample to predict movement profiles. The predicted movement profile was compared with original behavioral variables to compute the proportion of explained deviance/variance in representation of the model performance. We repeated this procedure 100 times and averaged the explained deviance/variance as cross-validated GLM performance for the session. The proportion of deviance explained in Poisson models is a similar parameter as the variance explained in a Gaussian model (R^2) to measure model performance.

Decoding waterspout locations with spikes

Naive Bayes classifiers were used to predict either five waterspout locations directly or three waterspout location classes relative to bodyside [106]. For decoding across behavioral phases, spiking activity during six behavioral windows from individual trials within a session was used. The PCA decomposed activity of all simultaneously recorded neurons from the same session. Decoding with smoothed original spikes result in same conclusions. The six behavioral windows for each trial are as follows: 1) water delivery, the 100 msec window after pump onset; 2) pre-lift, the 200 msec window before lift; 3) reach, the window from lift to advance endpoint; 4) grasp/withdraw-to-drink, the window from advance endpoint to the first consumption lick; 5) drink, the window from the first consumption lick to the last consumption lick in the trial. Within each behavioral window, the average of all activities was used to obtain a single value representing the activity for that trial.

To decode the population activity of different neuron types, PCA decomposed and trial-averaged activity of randomly selected neurons from each neuron type. An equal number of neurons (sixteen neurons) was selected from each neuron type to control the impact of neuron number on classifier performance. The random neuron selection process was iterated 1000 times. During each neuron selection process, a 10-fold cross-validated prediction accuracy was computed. Briefly, classifiers were trained on 75% of the activity samples for each waterspout location to predict waterspout locations and then tested on the left-out samples (25%) for cross-validation.

Population neural trajectory

Concatenation was used to visualize and quantify time-evolution of trial-averaged population neural activity for different waterspout locations. Briefly, a data matrix of neural activity X size $n \times ct$ was compiled where n = number of neurons, c = the number of waterspout locations, and t = the number of time points (at 20 msec bins). This matrix contained the firing rates of every neuron for every condition and every analyzed time point [67]. Principal component analysis (PCA) was used to decompose the centered and normalized X from $n \times ct$ to $m \times ct$ with m orthogonal dimensions with MATLAB function *pca*. The normalized X were projected to the m orthogonal dimensions to get *score* ($m \times ct$) values for all time points along all dimensions. To visualize the time-evolution of population activity, we plotted the projection of the normalized X along the first three principal component.

Geometric properties of population neural trajectories

To quantify the properties of population neural trajectories and eliminate the impact of neuron number on comparisons among different neuron types, an equal number of neurons (sixteen neurons) was randomly selected from each neuron type for PCA. To estimate geometric metrics of the population neural trajectories, the random resampling process was repeated 1000 times for all given measurements. 1) The smoothness of a neural trajectory was represented by summed jerk, which is the third time derivative of PCA score. The jerk of different orthogonal PCA dimensions was Euclidean normalized and summed. Lower common logarithm of the summed jerk indicated higher smoothness of a trajectory. 2) The dissimilarity in geometric shapes between a pair of neural trajectories was computed using the best shape-preserving Euclidean transformation between the two trajectories using the MATLAB *procrustes* function. Procrustes distance was quantified for each pair of neural trajectories at different waterspout locations. Lower values of Procrustes distance reflected higher similarity in shapes. 3) To characterize the spatial closeness between a pair of neural trajectories across different time points, we characterized the response distance between each pair of different neural trajectories using Euclidean distance along orthogonal PCA dimensions. Lower response distances between two samples were computed to represent similar neural activity.

SUPPLEMENTAL INFORMATION

Supplementary video 1

Network tracked behavior videos, 3D-reconstructed hand trajectories, and 4 moment-by-moment behavior profiles across P1, P3 and P5 waterspout locations are shown. Movement segments including lift, aim, advance, withdraw/grasp, and drink were annotated based on automatically identified key behavioral frames and labeled with different colors. Transitional frames were labeled with makers of different shapes. Behavioral profiles include hand to target distance, moving speed, supination score and waterspout aiming score of the left reaching hand. Random example trials from five waterspout locations were arranged sequentially. The movie is played at 1/40 of the original behavior speed.

Supplementary video 2

Cortex-wide calcium dynamics during reach and grasp to drink behavior at fixed waterspout location P2 and at five different waterspout locations for different projection neuron types. Activity was aligned at waterspout contact (0 msec) and averaged across all trials in each example session. Video is at the original speed.

Supplementary video 3

Photoinhibition of contralateral MOs-c of *Pvalb-ires-cre;Ai32* mouse after animal lifted the hand from an example session. The turquoise color square indicates inhibition light on. In the first part of video, the mouse failed to grasp the target upon inhibition. In the second part, the animal hit the target and grasped the water but failed to supinate and consume from the hand upon inhibition. Video is at 1/10 of the original behavior speed.

Supplementary video 4

Photoinhibition of other cortical nodes (SSp-ul, Prt, MO-orf) during reaching of a *Pvalb-ires-cre;Ai32* mouse. The turquoise color square indicates inhibition light on. Hand movement trajectories are superimposed on the video. Video is at 1/10 of the original speed.

Supplementary video 5

Population neural activity. The first part shows population neural trajectories from an example session with Gaussian-process factor analysis (FA). Each line with light color represents a trial. Thicker lines with dark color are the average across all trials. Lift, advance endpoint, and first hand lick time points are indicated by triangles, circles, and crosses, respectively. Line colors correspond to five different waterspout locations. The following parts show trial-averaged and session-pooled population neural activity with principal component (PC) analysis for all recorded neurons and for different projection neuron types. See Methods for details of the computation of population neural trajectory.

Figure 1

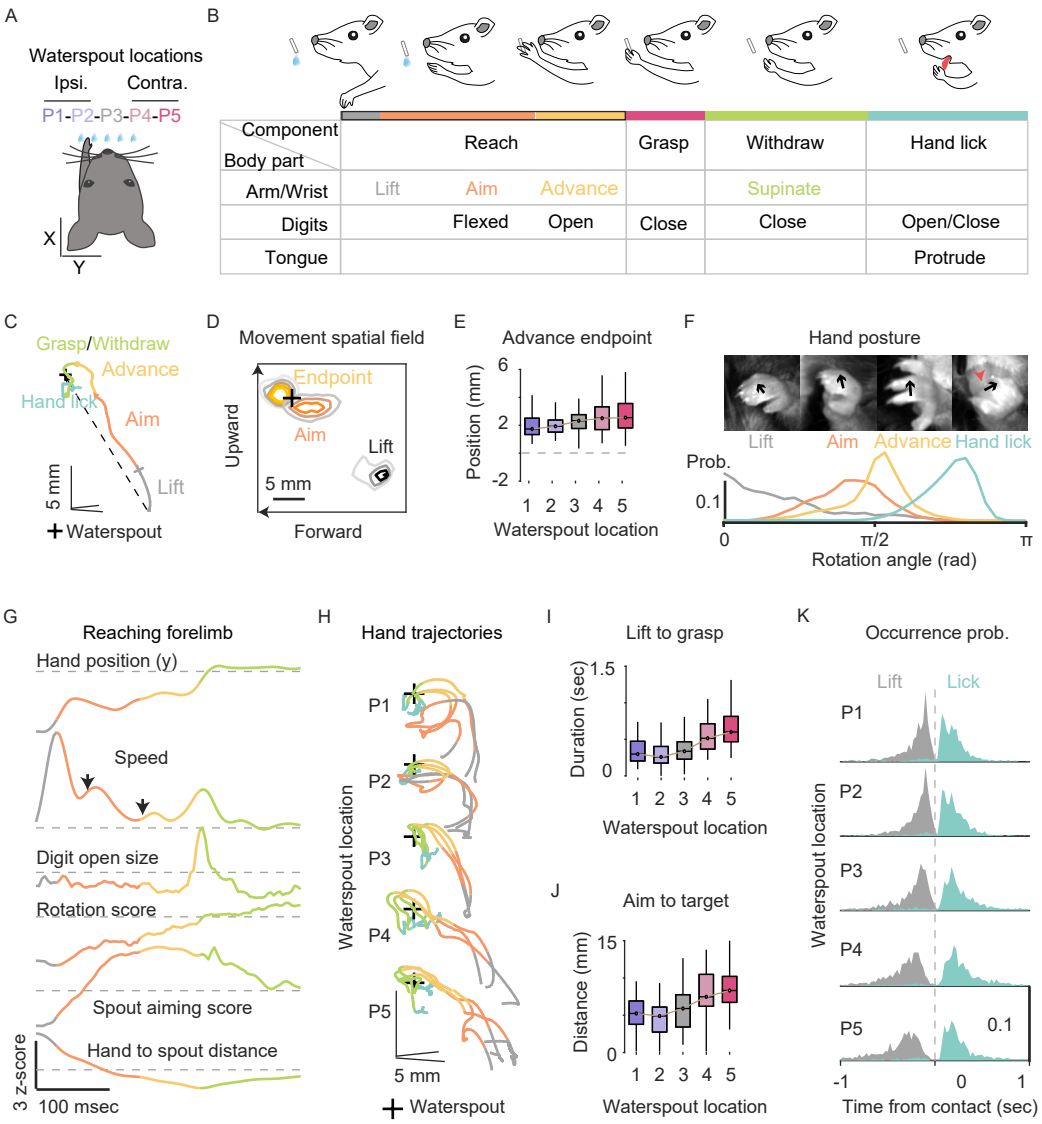

A
Waterspout locations
Ipsi.
Contra.
P1-P2-P3-P4-P5
X
Y
B
Component
Body part
Reach
Grasp
Withdraw
Hand lick
Arm/Wrist
Lift
Aim
Advance
Supinate
Digits
Flexed
Open
Close
Close
Open/Close
Tongue
Protrude
C
Grasp/Withdraw
Advance
Hand lick
Aim
Lift
5 mm
Waterspout
D
Movement spatial field
Endpoint
Aim
Lift
Upward
5 mm
Forward
E
Advance endpoint
Position (mm)
6
2
-2
1 2 3 4 5
Waterspout location
F
Hand posture
Lift
Aim
Advance
Hand lick
Prob.
0.1
0
π/2
π
Rotation angle (rad)
G
Reaching forelimb
Hand position (y)
Speed
Digit open size
Rotation score
Spout aiming score
Hand to spout distance
3 z-score
100 msec
H
Hand trajectories
P1
P2
P3
P4
P5
Waterspout location
5 mm
Waterspout
I
Lift to grasp
Duration (sec)
1.5
0
1 2 3 4 5
Waterspout location
J
Aim to target
Distance (mm)
15
0
1 2 3 4 5
Waterspout location
K
Occurrence prob.
Lift
Lick
P1
P2
P3
P4
P5
0.1
Waterspout location
-1
0
1
Time from contact (sec)

Figure 2

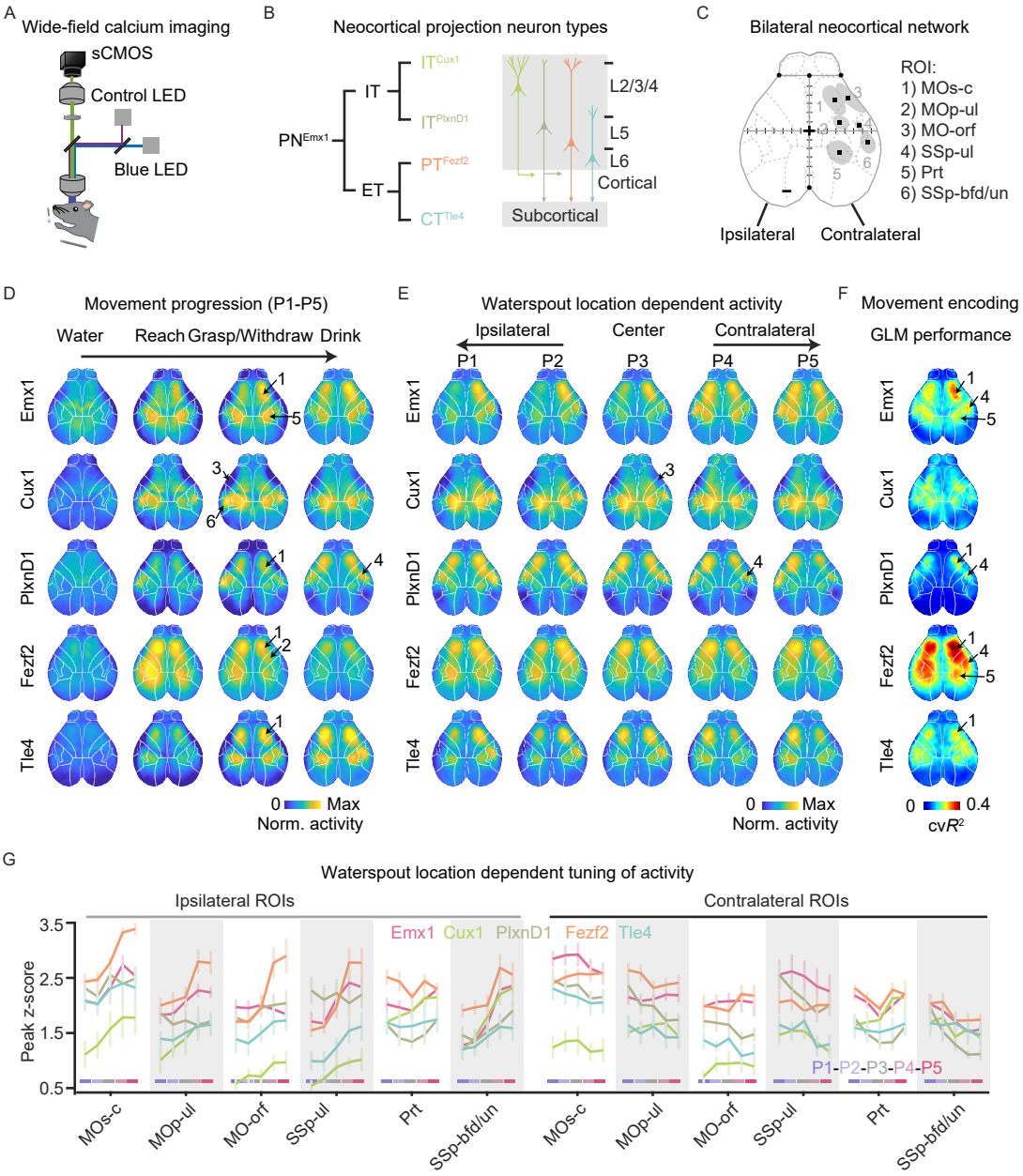

A
Wide-field calcium imaging
sCMOS
Control LED
Blue LED
B
Neocortical projection neuron types
PN^Emx1
IT
ET
IT^Cux1
IT^PlxnD1
PT^Fezf2
CT^Tle4
L2/3/4
L5
L6
Cortical
Subcortical
C
Bilateral neocortical network
ROI:
1) MOs-c
2) MOp-ul
3) MO-orf
4) SSp-ul
5) Prt
6) SSp-bfd/un
Ipsilateral
Contralateral
D
Movement progression (P1-P5)
Water
Reach
Grasp/Withdraw
Drink
Emx1
Cux1
PlxnD1
Fezf2
Tle4
0
Max
Norm. activity
E
Waterspout location dependent activity
Ipsilateral
Center
Contralateral
P1
P2
P3
P4
P5
0
Max
Norm. activity
F
Movement encoding
GLM performance
0
0.4
cvR²
G
Waterspout location dependent tuning of activity
Ipsilateral ROIs
Contralateral ROIs
Emx1
Cux1
PlxnD1
Fezf2
Tle4
Peak z-score
3.5
2.5
1.5
0.5
P1-P2-P3-P4-P5
MOs-c
MOp-ul
MO-orf
SSp-ul
Prt
SSp-bfd/un

Figure 3

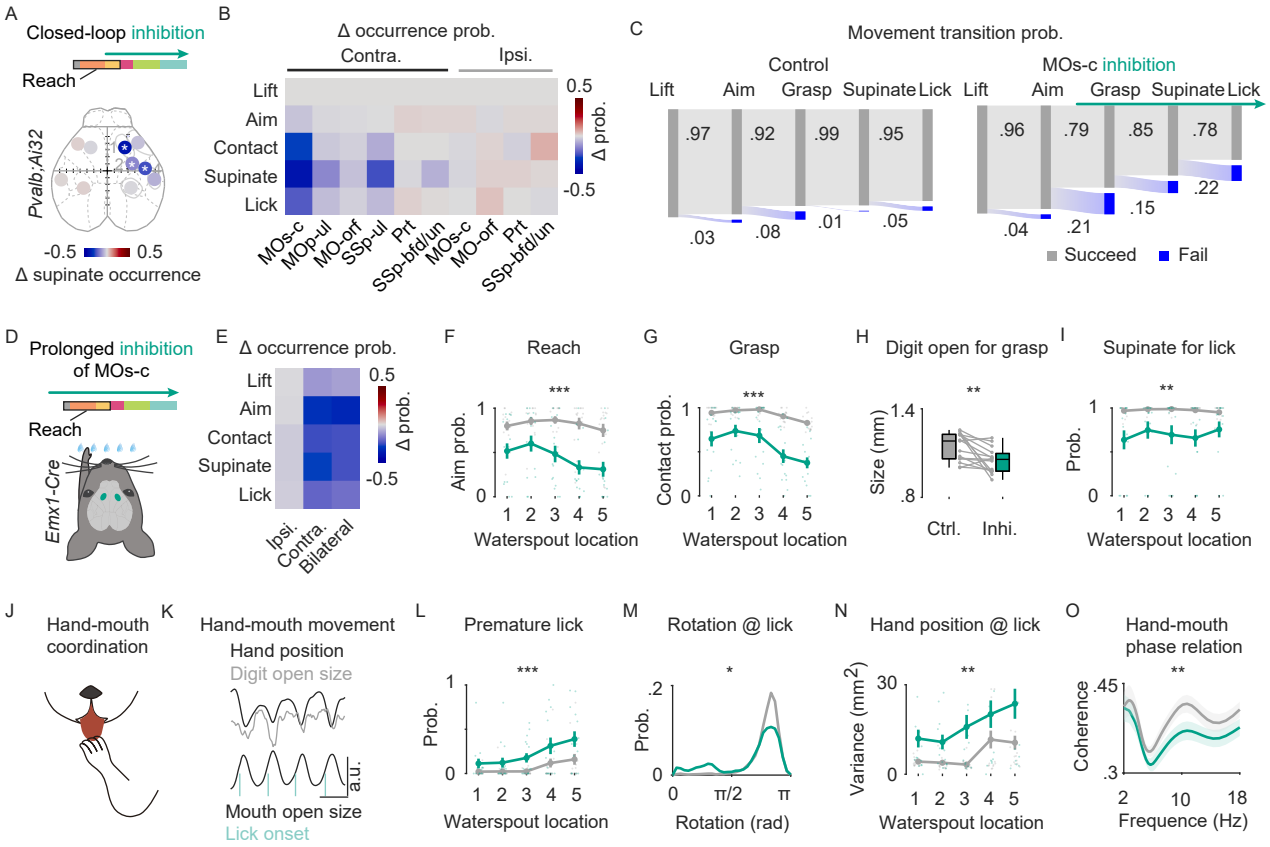

A
Closed-loop inhibition
Reach
Pvalb;Ai32
-0.5
0.5
Δ supinate occurrence
B
Δ occurrence prob.
Contra.
Ipsi.
Lift
Aim
Contact
Supinate
Lick
MOs-c
MOp-ul
MO-orf
SSp-ul
Prt
SSp-bfd/un
MOs-c
MO-orf
Prt
SSp-bfd/un
0.5
Δ prob.
-0.5
C
Movement transition prob.
Control
MOs-c inhibition
Lift
Aim
Grasp
Supinate
Lick
.97
.92
.99
.95
.03
.08
.01
.05
.96
.79
.85
.78
.04
.21
.15
.22
Succeed
Fail
D
Prolonged inhibition of MOs-c
Reach
Emx1-Cre
E
Δ occurrence prob.
Lift
Aim
Contact
Supinate
Lick
Ipsi.
Contra.
Bilateral
0.5
Δ prob.
-0.5
F
Reach
***
Aim prob.
1
0
1 2 3 4 5
Waterspout location
G
Grasp
***
Contact prob.
1
0
1 2 3 4 5
Waterspout location
H
Digit open for grasp
**
Size (mm)
1.4
.8
Ctrl.
Inhi.
I
Supinate for lick
**
Prob.
1
0
1 2 3 4 5
Waterspout location
J
Hand-mouth coordination
K
Hand-mouth movement
Hand position
Digit open size
a.u.
Mouth open size
Lick onset
L
Premature lick
***
Prob.
1
0
1 2 3 4 5
Waterspout location
M
Rotation @ lick
*
Prob.
.2
0
0
π/2
π
Rotation (rad)
N
Hand position @ lick
**
Variance (mm2)
30
0
1 2 3 4 5
Waterspout location
O
Hand-mouth phase relation
**
Coherence
.45
.3
2
10
18
Frequence (Hz)

Figure 4

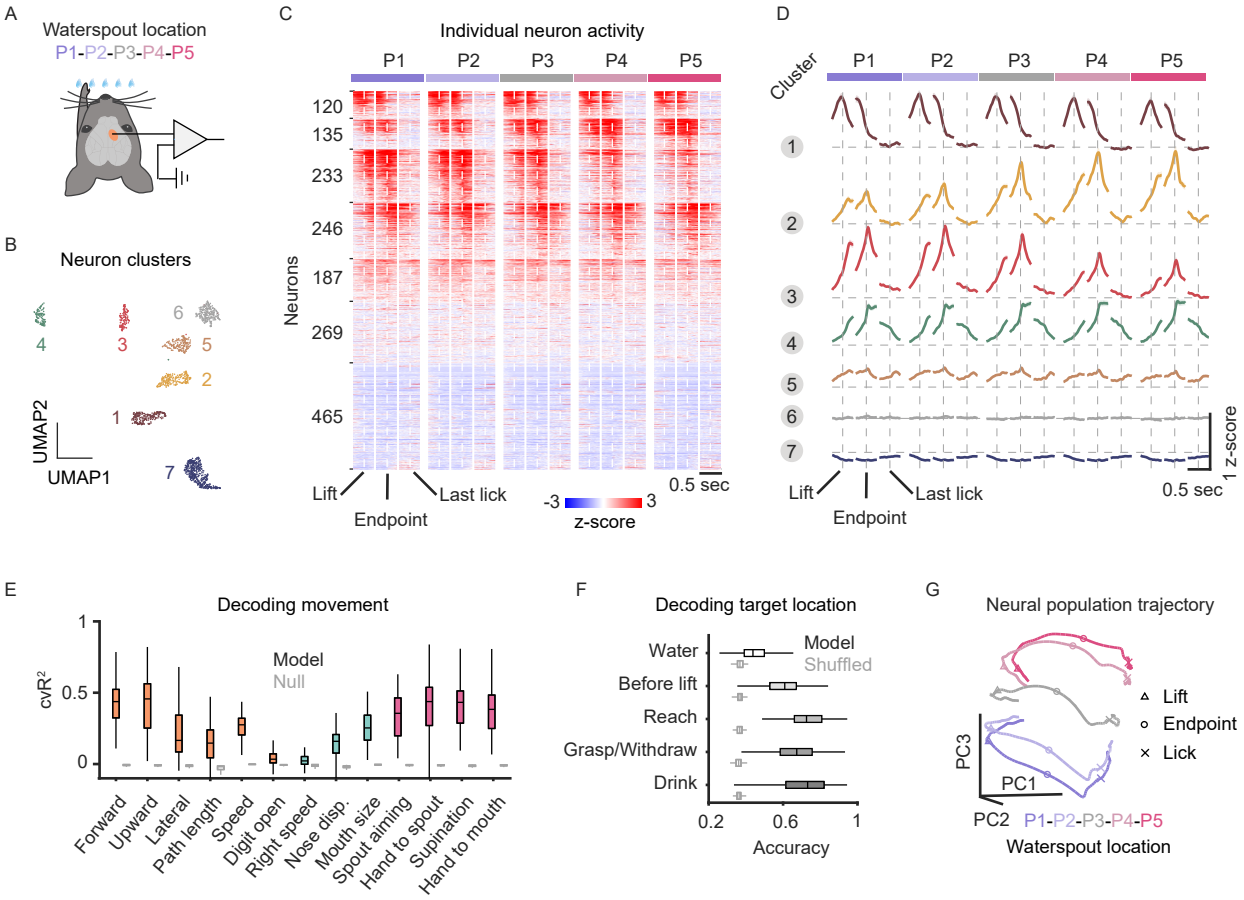

A
Waterspout location
P1-P2-P3-P4-P5
B
Neuron clusters
UMAP2
UMAP1
C
Individual neuron activity
P1
P2
P3
P4
P5
Neurons
120
135
233
246
187
269
465
Lift
Endpoint
Last lick
-3
3
z-score
0.5 sec
D
Cluster
1
2
3
4
5
6
7
1 z-score
Lift
Endpoint
Last lick
0.5 sec
E
Decoding movement
cvR²
Model
Null
Forward
Upward
Lateral
Path length
Speed
Digit open
Right speed
Nose disp.
Mouth size
Spout aiming
Hand to spout
Supination
Hand to mouth
F
Decoding target location
Water
Before lift
Reach
Grasp/Withdraw
Drink
Model
Shuffled
Accuracy
G
Neural population trajectory
Lift
Endpoint
Lick
PC3
PC1
PC2
P1-P2-P3-P4-P5
Waterspout location

Figure 5

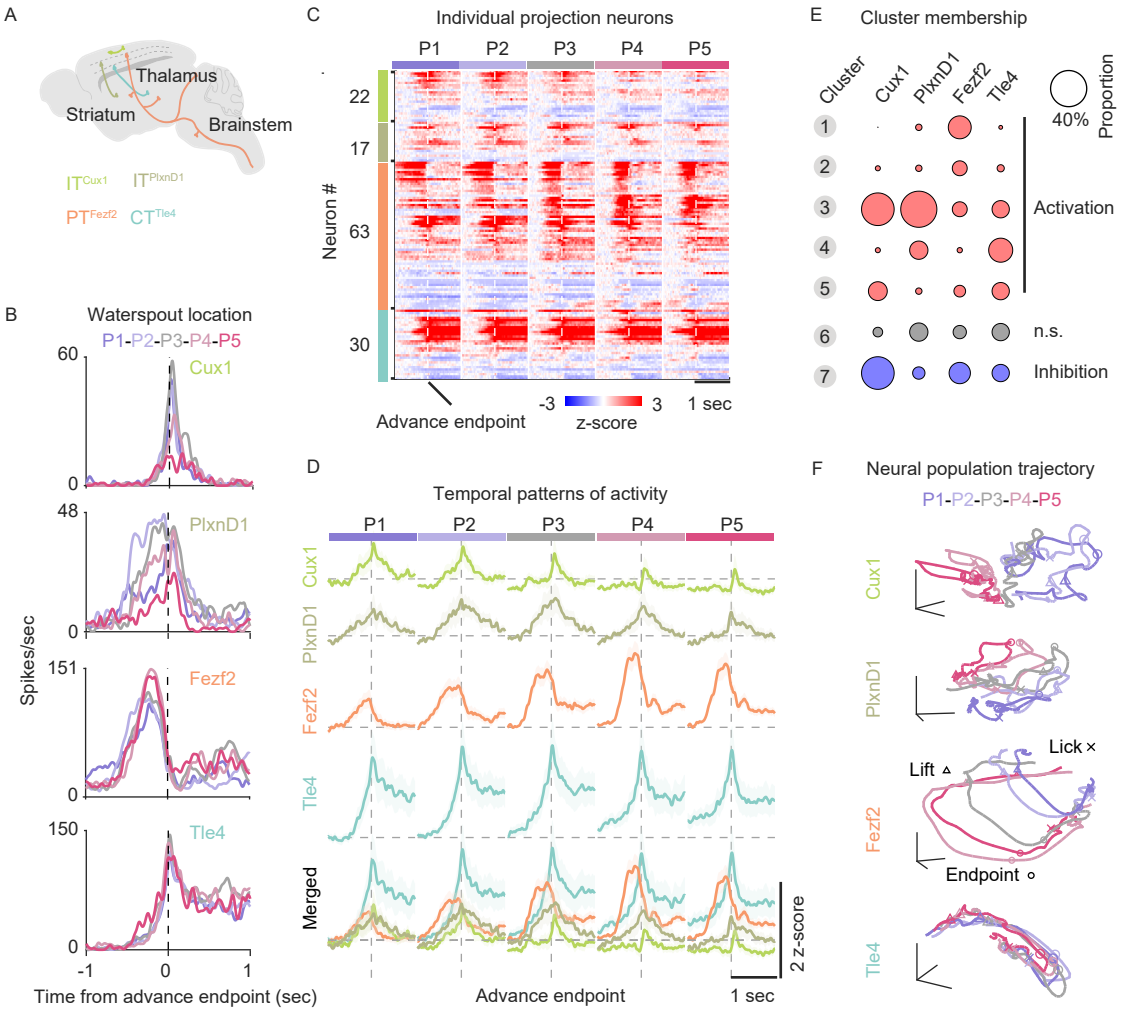

A
Thalamus
Striatum
Brainstem
IT Cux1
IT PlxnD1
PT Fezf2
CT Tle4
B
Waterspout location
P1-P2-P3-P4-P5
Cux1
PlxnD1
Fezf2
Tle4
Spikes/sec
Time from advance endpoint (sec)
C
Individual projection neurons
P1
P2
P3
P4
P5
Neuron #
Advance endpoint
z-score
1 sec
D
Temporal patterns of activity
Merged
2 z-score
Advance endpoint
1 sec
E
Cluster membership
Cluster
Proportion
40%
Activation
n.s.
Inhibition
F
Neural population trajectory
P1-P2-P3-P4-P5
Lift
Lick
Endpoint

Figure 6

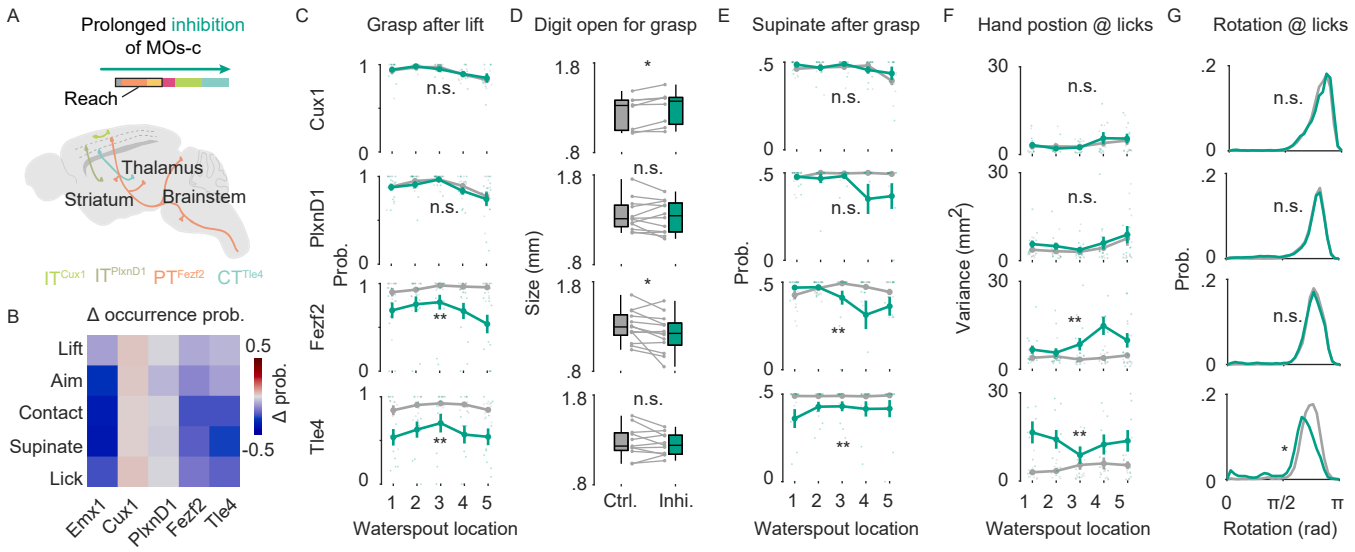

A
Prolonged inhibition of MOs-c
Reach
Thalamus
Striatum
Brainstem
IT Cux1
IT PlxnD1
PT Fezf2
CT Tle4
B
Δ occurrence prob.
Lift
Aim
Contact
Supinate
Lick
Emx1
Cux1
PlxnD1
Fezf2
Tle4
0.5
-0.5
Δ prob.
C
Grasp after lift
Cux1
PlxnD1
Fezf2
Tle4
Prob.
n.s.
**
1 2 3 4 5
Waterspout location
D
Digit open for grasp
Size (mm)
1.8
.8
*
n.s.
Ctrl.
Inhi.
E
Supinate after grasp
Prob.
.5
0
n.s.
**
1 2 3 4 5
Waterspout location
F
Hand postion @ licks
Variance (mm²)
30
0
n.s.
**
1 2 3 4 5
Waterspout location
G
Rotation @ licks
Prob.
.2
0
n.s.
*
0
π/2
π
Rotation (rad)

Figure S1

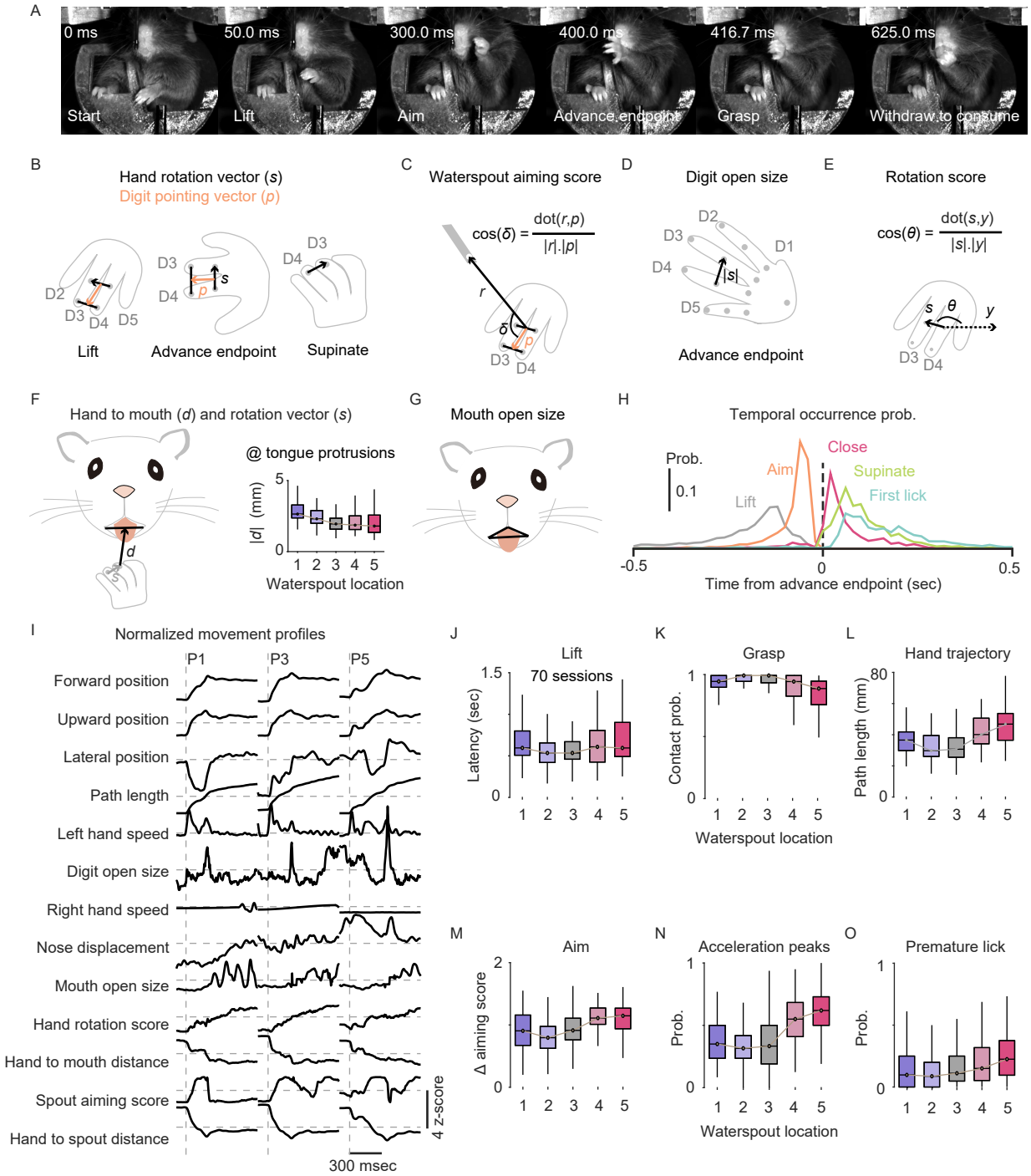

A
0 ms
50.0 ms
300.0 ms
400.0 ms
416.7 ms
625.0 ms
Start
Lift
Aim
Advance endpoint
Grasp
Withdraw to consume
B
Hand rotation vector (s)
Digit pointing vector (p)
Lift
Advance endpoint
Supinate
C
Waterspout aiming score
cos(δ) = dot(r,p) / |r|.|p|
D
Digit open size
Advance endpoint
E
Rotation score
cos(θ) = dot(s,y) / |s|.|y|
F
Hand to mouth (d) and rotation vector (s)
@ tongue protrusions
|d| (mm)
Waterspout location
G
Mouth open size
H
Temporal occurrence prob.
Prob.
0.1
Lift
Aim
Close
Supinate
First lick
-0.5
0
0.5
Time from advance endpoint (sec)
I
Normalized movement profiles
P1
P3
P5
Forward position
Upward position
Lateral position
Path length
Left hand speed
Digit open size
Right hand speed
Nose displacement
Mouth open size
Hand rotation score
Hand to mouth distance
Spout aiming score
Hand to spout distance
4 z-score
300 msec
J
Lift
70 sessions
Latency (sec)
K
Grasp
Contact prob.
Waterspout location
L
Hand trajectory
Path length (mm)
M
Aim
Δ aiming score
N
Acceleration peaks
Prob.
Waterspout location
O
Premature lick
Prob.

Figure S2

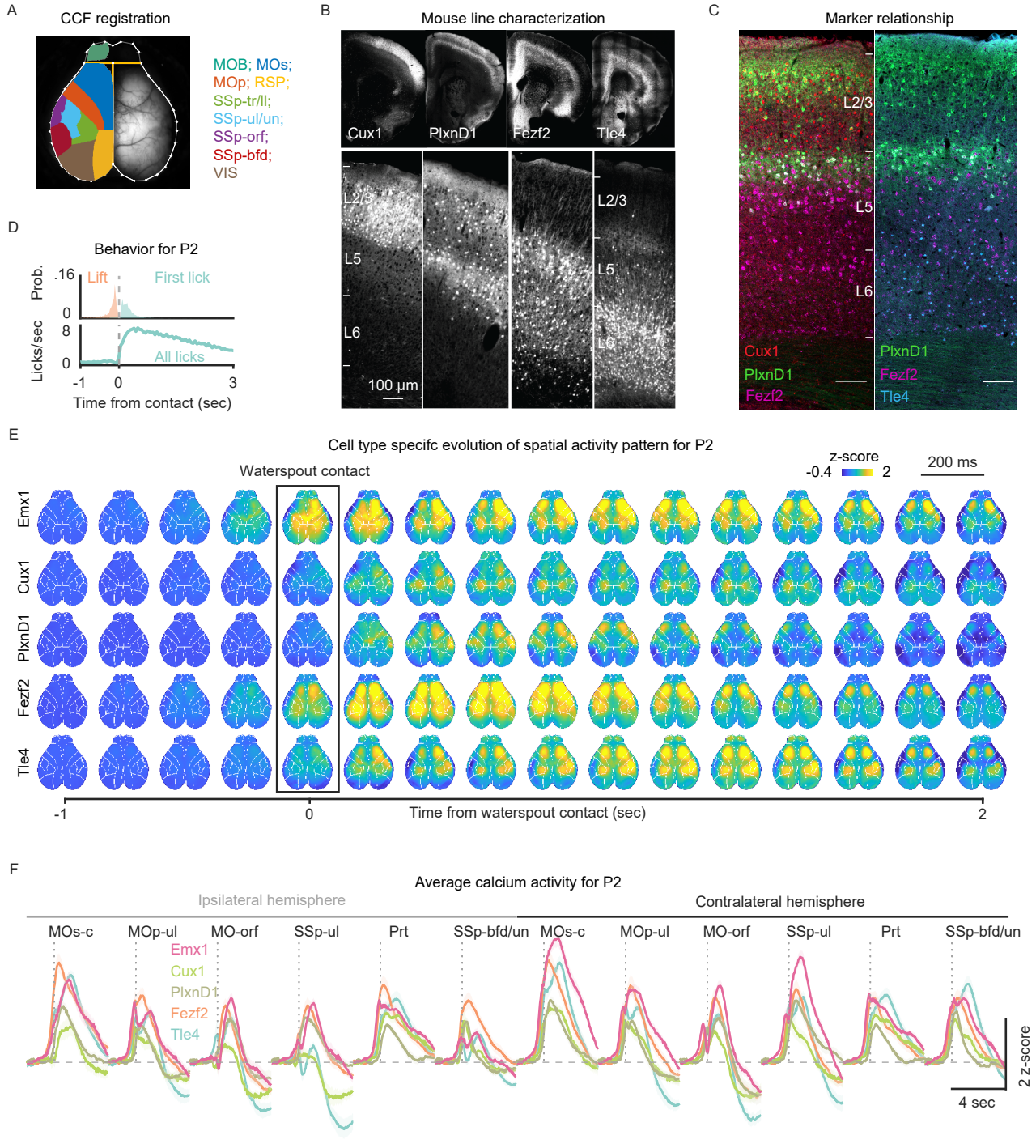

A
CCF registration
MOB; MOs;
MOp; RSP;
SSp-tr/ll;
SSp-ul/un;
SSp-orf;
SSp-bfd;
VIS
B
Mouse line characterization
Cux1
PlxnD1
Fezf2
Tle4
L2/3
L5
L6
100 µm
C
Marker relationship
L2/3
L5
L6
Cux1
PlxnD1
Fezf2
PlxnD1
Fezf2
Tle4
D
Behavior for P2
Prob.
.16
Lift
First lick
0
8
Licks/sec
0
All licks
-1
0
3
Time from contact (sec)
E
Cell type specifc evolution of spatial activity pattern for P2
Waterspout contact
z-score
-0.4
2
200 ms
Emx1
Cux1
PlxnD1
Fezf2
Tle4
-1
0
Time from waterspout contact (sec)
2
F
Average calcium activity for P2
Ipsilateral hemisphere
Contralateral hemisphere
MOs-c
MOp-ul
MO-orf
SSp-ul
Prt
SSp-bfd/un
MOs-c
MOp-ul
MO-orf
SSp-ul
Prt
SSp-bfd/un
Emx1
Cux1
PlxnD1
Fezf2
Tle4
2 z-score
4 sec

Figure S3

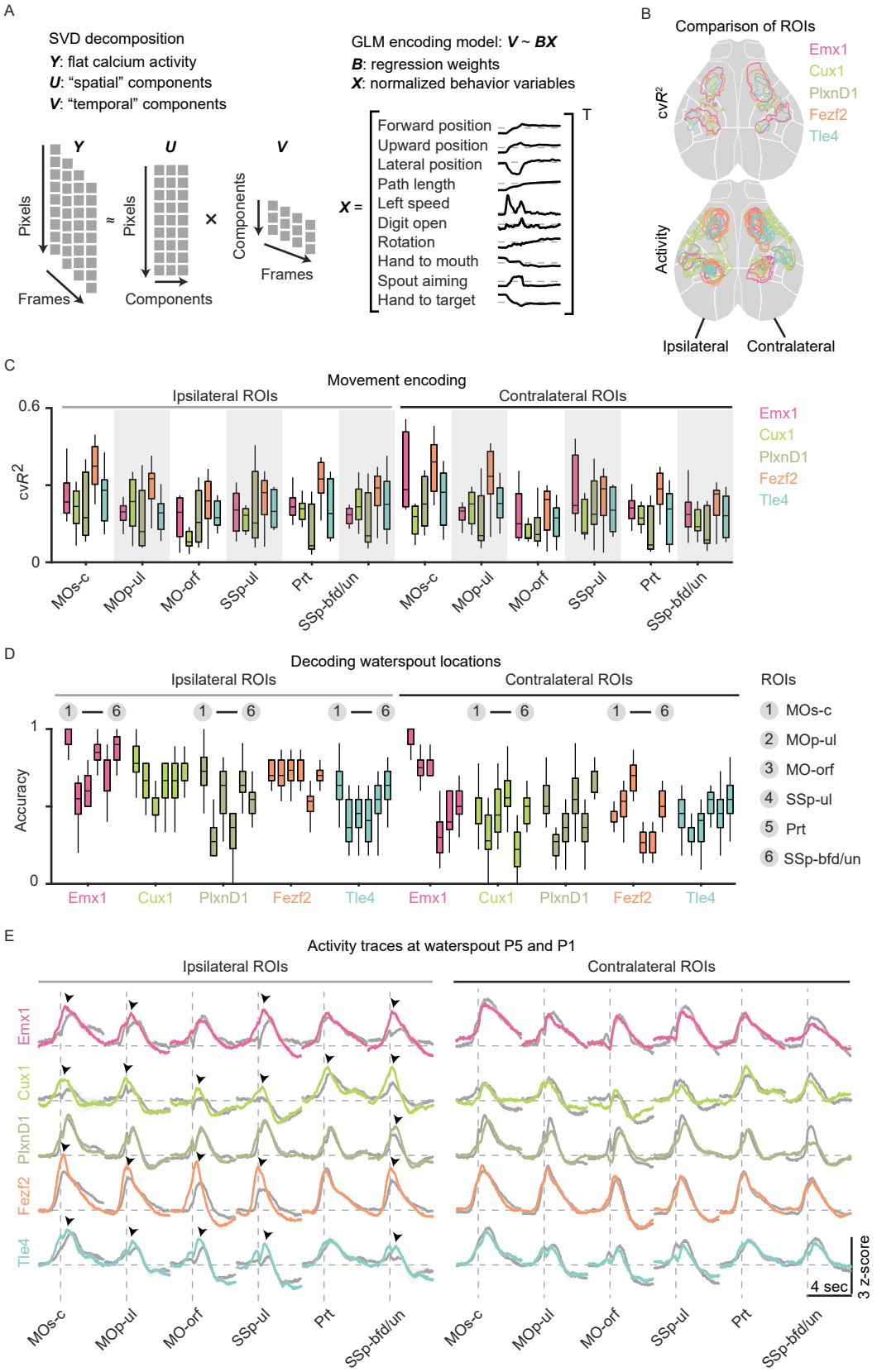

A
SVD decomposition
Y: flat calcium activity
U: "spatial" components
V: "temporal" components
Y
U
V
Pixels
Frames
≈
Pixels
Components
×
Components
Frames
GLM encoding model: V ~ BX
B: regression weights
X: normalized behavior variables
X =
T
Forward position
Upward position
Lateral position
Path length
Left speed
Digit open
Rotation
Hand to mouth
Spout aiming
Hand to target
B
Comparison of ROIs
Emx1
Cux1
PlxnD1
Fezf2
Tle4
cvR²
Activity
Ipsilateral
Contralateral
C
Movement encoding
Ipsilateral ROIs
Contralateral ROIs
0.6
cvR2
0
MOs-c
MOp-ul
MO-orf
SSp-ul
Prt
SSp-bfd/un
Emx1
Cux1
PlxnD1
Fezf2
Tle4
D
Decoding waterspout locations
Ipsilateral ROIs
Contralateral ROIs
1 — 6
Accuracy
1
0
Emx1
Cux1
PlxnD1
Fezf2
Tle4
ROIs
1 MOs-c
2 MOp-ul
3 MO-orf
4 SSp-ul
5 Prt
6 SSp-bfd/un
E
Activity traces at waterspout P5 and P1
Ipsilateral ROIs
Contralateral ROIs
Emx1
Cux1
PlxnD1
Fezf2
Tle4
4 sec
3 z-score
MOs-c
MOp-ul
MO-orf
SSp-ul
Prt
SSp-bfd/un

Figure S4

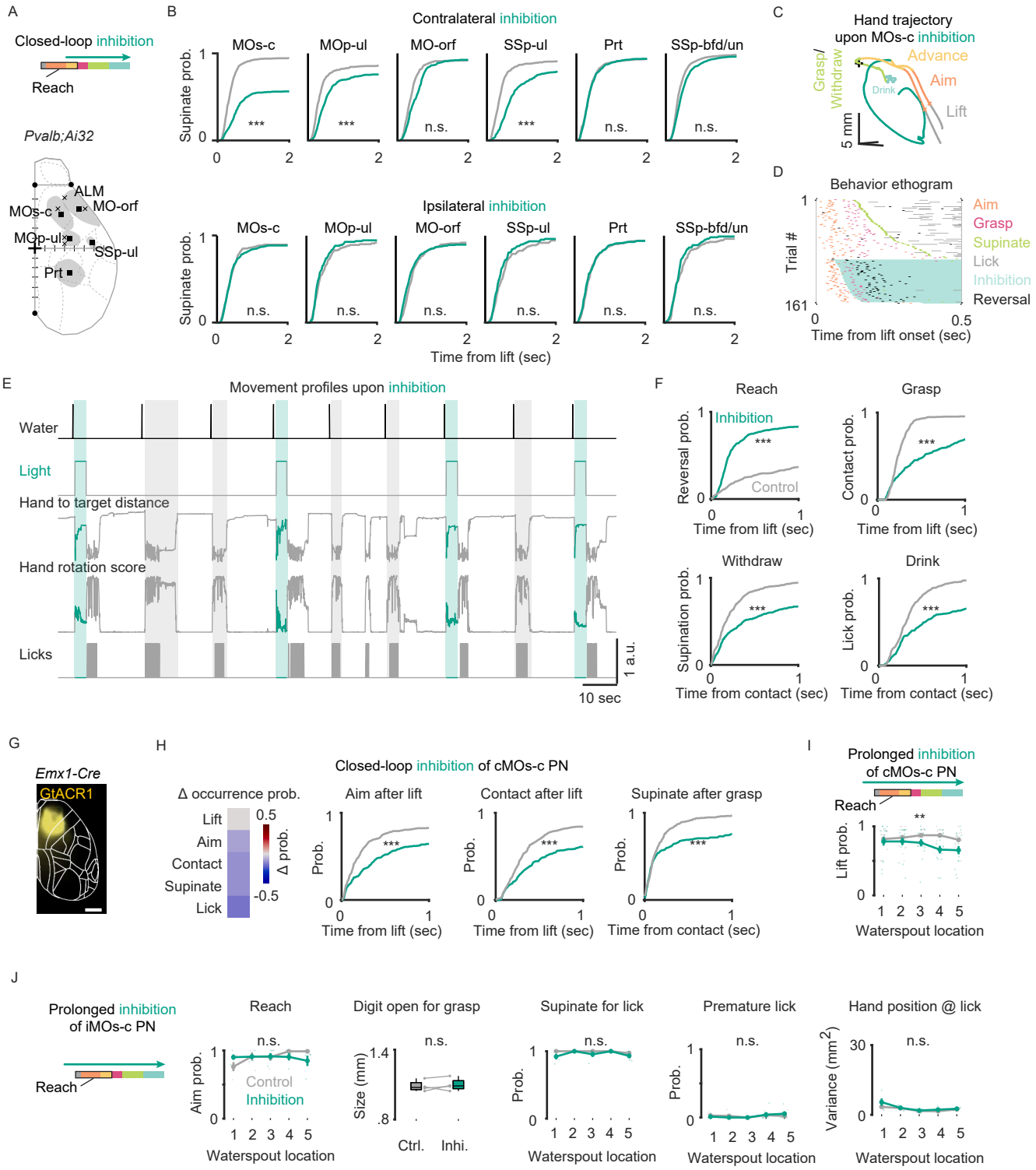

A
Closed-loop inhibition
Reach
Pvalb;Ai32
ALM
MOs-c
MO-orf
MOp-ul
SSp-ul
Prt
B
Contralateral inhibition
MOs-c
MOp-ul
MO-orf
SSp-ul
Prt
SSp-bfd/un
Supinate prob.
***
n.s.
Ipsilateral inhibition
Time from lift (sec)
C
Hand trajectory
upon MOs-c inhibition
Grasp/
Withdraw
Advance
Aim
Drink
Lift
5 mm
D
Behavior ethogram
Trial #
Aim
Grasp
Supinate
Lick
Inhibition
Reversal
Time from lift onset (sec)
E
Movement profiles upon inhibition
Water
Light
Hand to target distance
Hand rotation score
Licks
1 a.u.
10 sec
F
Reach
Grasp
Inhibition
Control
Reversal prob.
Contact prob.
Time from lift (sec)
Withdraw
Drink
Supination prob.
Lick prob.
Time from contact (sec)
G
Emx1-Cre
GtACR1
H
Closed-loop inhibition of cMOs-c PN
Δ occurrence prob.
Lift
Aim
Contact
Supinate
Lick
0.5
-0.5
Δ prob.
Aim after lift
Contact after lift
Supinate after grasp
Prob.
Time from lift (sec)
Time from contact (sec)
I
Prolonged inhibition
of cMOs-c PN
Reach
Lift prob.
**
Waterspout location
J
Prolonged inhibition
of iMOs-c PN
Reach
Reach
Digit open for grasp
Supinate for lick
Premature lick
Hand position @ lick
Aim prob.
Control
Inhibition
Size (mm)
Ctrl.
Inhi.
Prob.
Variance (mm2)
Waterspout location

Figure S5

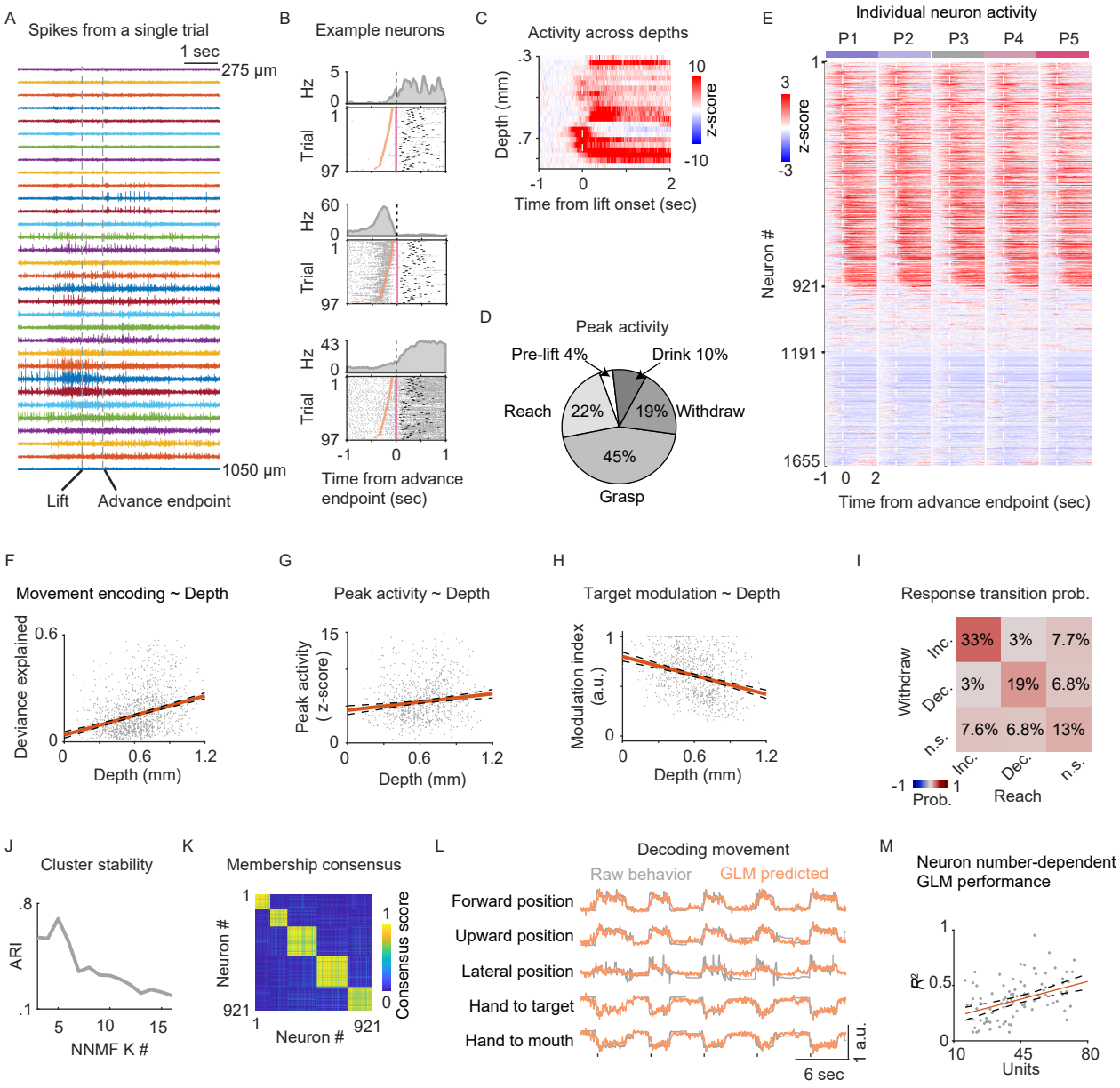

A
Spikes from a single trial
1 sec
275 µm
1050 µm
Lift
Advance endpoint
B
Example neurons
Hz
Trial
Time from advance endpoint (sec)
C
Activity across depths
Depth (mm)
z-score
Time from lift onset (sec)
D
Peak activity
Pre-lift 4%
Drink 10%
Reach 22%
19% Withdraw
45%
Grasp
E
Individual neuron activity
P1
P2
P3
P4
P5
z-score
Neuron #
Time from advance endpoint (sec)
F
Movement encoding ~ Depth
Deviance explained
Depth (mm)
G
Peak activity ~ Depth
Peak activity (z-score)
Depth (mm)
H
Target modulation ~ Depth
Modulation index (a.u.)
Depth (mm)
I
Response transition prob.
Withdraw
Inc.
Dec.
n.s.
33%
3%
7.7%
3%
19%
6.8%
7.6%
6.8%
13%
Reach
Prob.
J
Cluster stability
ARI
NNMF K #
K
Membership consensus
Neuron #
Consensus score
L
Decoding movement
Raw behavior
GLM predicted
Forward position
Upward position
Lateral position
Hand to target
Hand to mouth
1 a.u.
6 sec
M
Neuron number-dependent GLM performance
Units

Figure S6

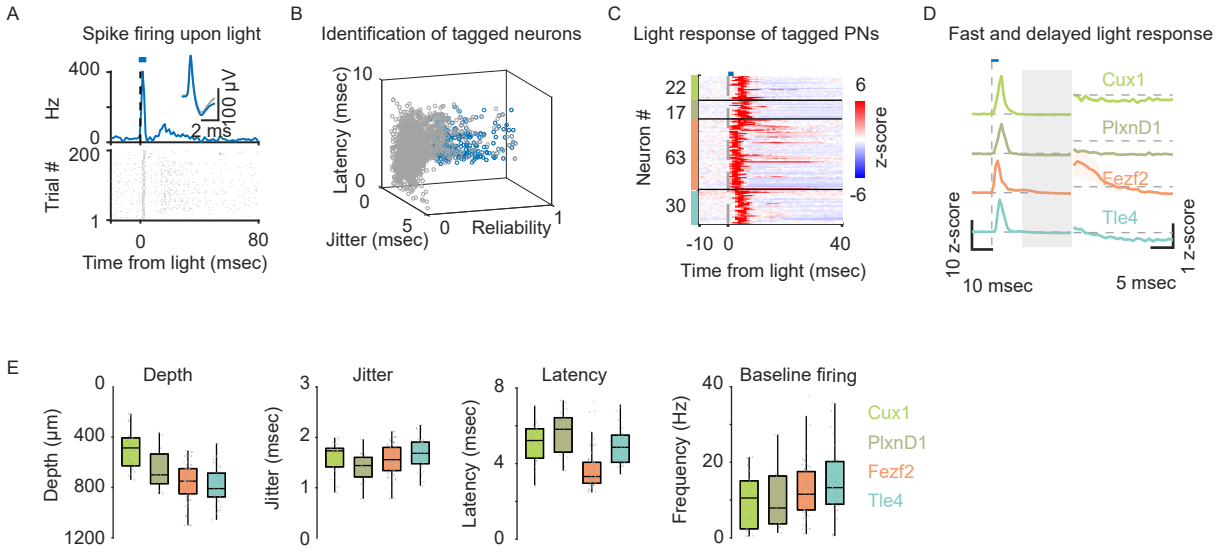

A
Spike firing upon light
400
Hz
0
200
Trial #
1
0
80
Time from light (msec)
100 µV
2 ms
B
Identification of tagged neurons
Latency (msec)
10
0
0
5
Jitter (msec)
0
Reliability
1
C
Light response of tagged PNs
22
17
63
30
Neuron #
-10
0
40
Time from light (msec)
6
-6
z-score
D
Fast and delayed light response
Cux1
PlxnD1
Fezf2
Tle4
10 z-score
1 z-score
10 msec
5 msec
E
Depth
Depth (µm)
0
400
800
1200
Jitter
Jitter (msec)
3
2
1
0
Latency
Latency (msec)
8
4
0
Baseline firing
Frequency (Hz)
40
20
0
Cux1
PlxnD1
Fezf2
Tle4

Figure S7

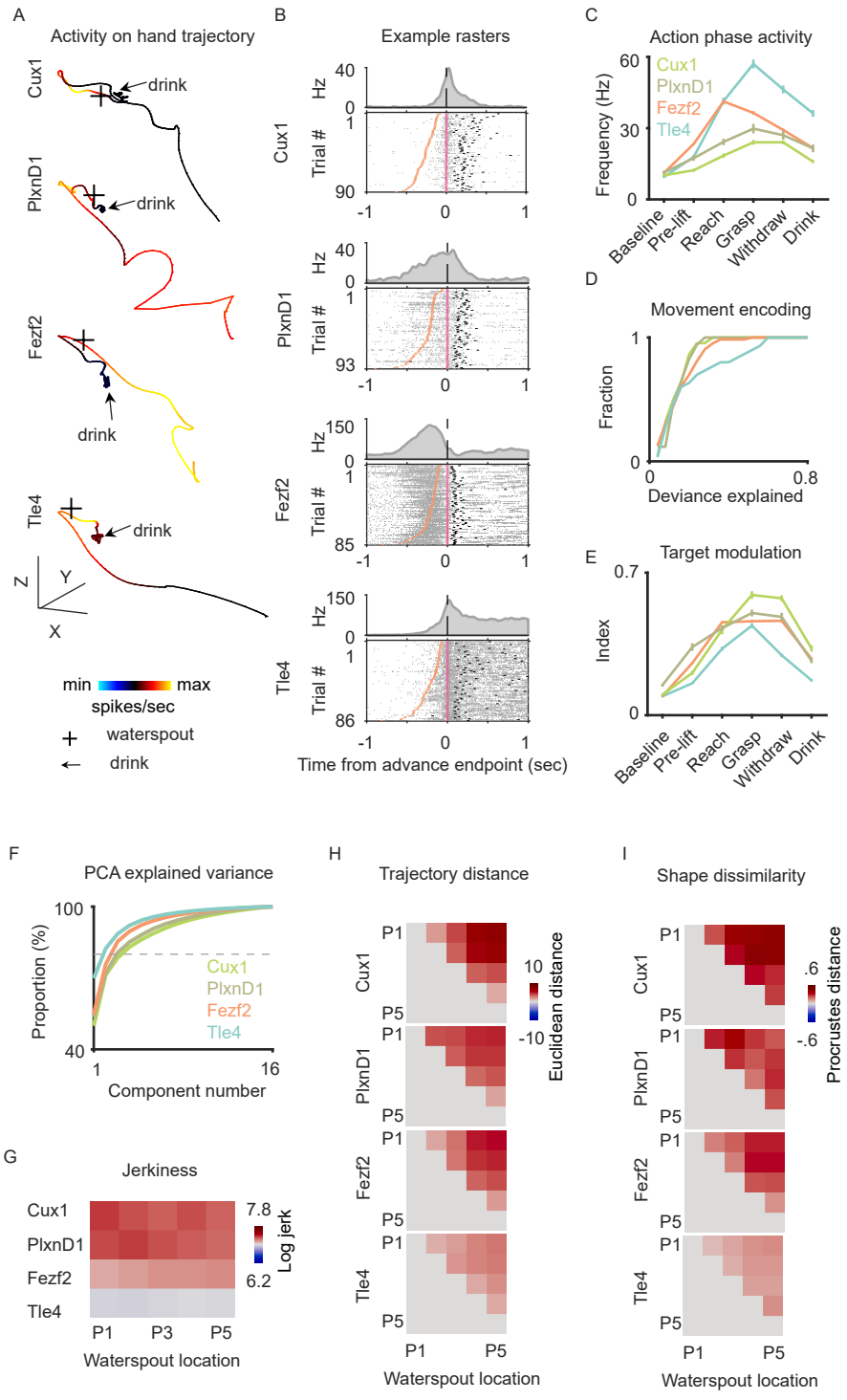
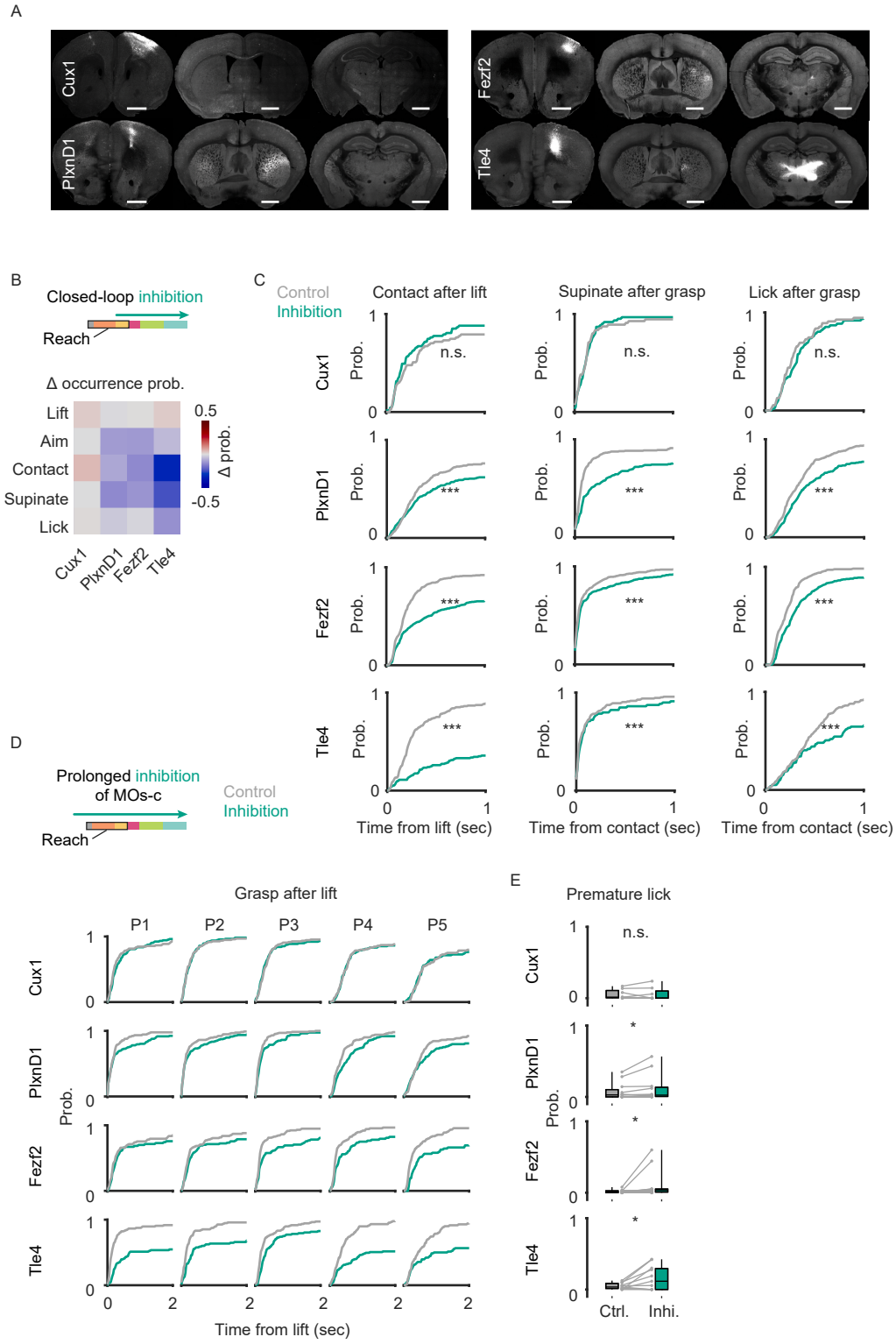

A
Activity on hand trajectory
Cux1
PlxnD1
Fezf2
Tle4
drink
Z
Y
X
min
max
spikes/sec
waterspout
drink
B
Example rasters
Hz
Trial #
Time from advance endpoint (sec)
C
Action phase activity
Frequency (Hz)
Cux1
PlxnD1
Fezf2
Tle4
Baseline
Pre-lift
Reach
Grasp
Withdraw
Drink
D
Movement encoding
Fraction
Deviance explained
E
Target modulation
Index
F
PCA explained variance
Proportion (%)
Component number
G
Jerkiness
Log jerk
Waterspout location
H
Trajectory distance
Euclidean distance
Waterspout location
I
Shape dissimilarity
Procrustes distance
Waterspout location

Figure S8


A
Cux1
PlxnD1
Fezf2
Tle4
B
Closed-loop inhibition
Reach
Δ occurrence prob.
Lift
Aim
Contact
Supinate
Lick
Cux1
PlxnD1
Fezf2
Tle4
0.5
-0.5
Δ prob.
C
Control
Inhibition
Contact after lift
Supinate after grasp
Lick after grasp
Prob.
n.s.
***
Time from lift (sec)
Time from contact (sec)
D
Prolonged inhibition of MOs-c
Reach
Control
Inhibition
Grasp after lift
P1
P2
P3
P4
P5
Time from lift (sec)
E
Premature lick
n.s.
*
Ctrl.
Inhi.