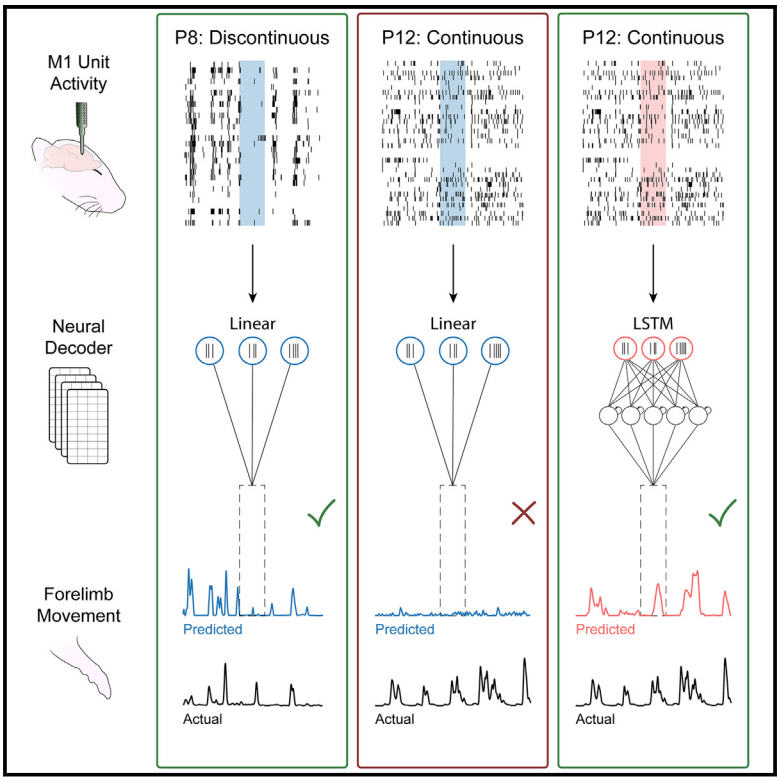
Report

Cell Reports

# Neural decoding reveals specialized kinematic tuning after an abrupt cortical transition

## Graphical abstract




## Authors

Ryan M. Glanz, Greta Sokoloff, Mark S. Blumberg

## Correspondence

mark-blumberg@uiowa.edu

## In brief

Cortical activity transitions from discontinuous to continuous around postnatal day 12 in rats. Glanz et al. use neural decoding to assess this transition's effect on sensory representations of limb kinematics in the primary motor cortex (M1). After the transition, M1 activity is more informationally dense, and its encoding scheme is more individualized.

## Highlights

- In P8 rats, a linear decoder reliably predicts limb movements from M1 activity
- At P12, M1 activity is continuous, and a nonlinear decoder is required
- M1 activity is more sensitive to "lesions" and encoding scheme "transplants" at P12
- P12 marks the onset of more sparse and individualized M1 movement representations

Glanz et al., 2023, Cell Reports *42*, 113119
September 26, 2023 
https://doi.org/10.1016/j.celrep.2023.113119

CellPress

Cell Reports

CellPress
OPEN ACCESS

Report

# Neural decoding reveals specialized kinematic tuning after an abrupt cortical transition

Ryan M. Glanz,[1] Greta Sokoloff,[1,2] and Mark S. Blumberg[1,2,3,*]
[1]Department of Psychological & Brain Sciences, University of Iowa, Iowa City, IA 52242, USA
[2]Iowa Neuroscience Institute, University of Iowa, Iowa City, IA 52242, USA
[3]Lead contact
*Correspondence: mark-blumberg@uiowa.edu
https://doi.org/10.1016/j.celrep.2023.113119

## SUMMARY

The primary motor cortex (M1) exhibits a protracted period of development, including the development of a sensory representation long before motor outflow emerges. In rats, this representation is present by postnatal day (P) 8, when M1 activity is "discontinuous." Here, we ask how the representation changes upon the transition to "continuous" activity at P12. We use neural decoding to predict forelimb movements from M1 activity and show that a linear decoder effectively predicts limb movements at P8 but not at P12; instead, a nonlinear decoder better predicts limb movements at P12. The altered decoder performance reflects increased complexity and uniqueness of kinematic information in M1. We next show that M1's representation at P12 is more susceptible to "lesioning" of inputs and "transplanting" of M1's encoding scheme from one pup to another. Thus, the emergence of continuous M1 activity signals the developmental onset of more complex, informationally sparse, and individualized sensory representations.

## INTRODUCTION

Perhaps the most striking feature of infant cortical activity is its discontinuity: periods of silence are punctuated by bursts of population-level activity.[1,2] This early phase of discontinuity ends with the sudden and dramatic onset of continuous cortical activity. Whereas in humans this transition occurs around the time of birth,[3] in infant rats and mice, it occurs at the end of the second postnatal week.[4–8] In fact, the onset of continuous cortical activity is but one aspect of a more global reorganization of brain dynamics that includes a shift in GABAergic functioning,[9,10] the proliferation and diversification of inhibitory interneurons,[11–14] accelerated myelin deposition,[15] and the onset of brain rhythms such as delta[16–18] and theta.[19]

In rats, continuous activity emerges in the primary motor cortex (M1) between postnatal days (P) 8 and P12.[6] Despite its name, M1 at these ages does not produce movement but instead functions exclusively as a somatosensory structure.[20–22] Accordingly, M1 activity reflects sensory input arising from self- or other-generated movements (i.e., reafference or exafference, respectively). Recently, we showed that neurons within M1's nascent somatosensory map are tuned to limb kinematics, as is the case for M1's adult motor map.[23,24] Specifically, we found precise sensory tuning to movement amplitude at P8—especially for the limb twitches that occur during active (rapid eye movement [REM]) sleep.[6] However, upon the emergence of continuous activity at P12, this tuning disappeared.

Was M1's kinematic tuning truly lost, or was it only obscured by the continuous activity? Here, we address this and related questions using neural decoding, a computational technique that allows us to predict the timing and amplitude of forelimb movements based solely on M1 activity. Our findings demonstrate that M1's processing of sensory information is intact but substantially transformed after this fundamental transition in cortical dynamics.

## RESULTS

We performed neural decoding using a previously published dataset from P8 and P12 rats (n = 8 at each age).[6] This dataset consists of 1-h recordings of M1 unit activity and video-based records of 3D forelimb displacement. For all decoding procedures, each 10-ms bin of forelimb displacement was predicted by M1 unit activity within a 240-ms window (i.e., 12 bins before and after). Each recording was divided into training (36 min; 60%) and testing (9 min; 15%) datasets; a validation dataset (15 min; 25%) was held back until the final model parameters were established. All analyses were performed on the validation dataset to ensure an unbiased assessment of decoder performance.[25]

### Movement encoding is obscured by continuous activity at P12

Discontinuous M1 activity presents a strong contrast between periods of movement (when reafference triggers neural activity) and rest (when there is little to no activity). Moreover, at P8, we previously found that forelimb twitches trigger rate-coded M1 responses that correlate with twitch amplitude.[6] Accordingly, we predicted that a linear decoder would accurately predict the temporal and spatial properties of forelimb movements at P8.

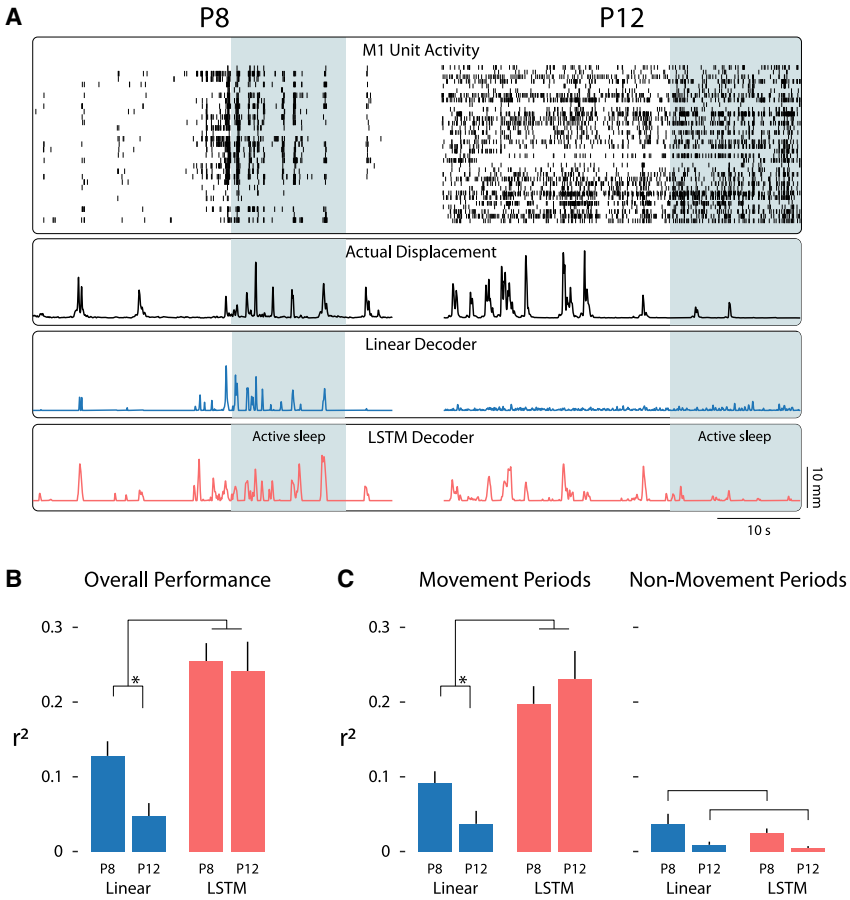



**Figure 1. Neural decoding in M1 predicts forelimb movements across the transition to continuous cortical activity**

(A) Representative data from a P8 (left) and a P12 (right) rat. From top to bottom: M1 unit activity where each row denotes an individual unit and each vertical tick denotes an action potential; actual forelimb displacement (black lines) in mm, representing the Euclidean distance traveled by the forelimb in 3D space; forelimb displacement as predicted by a linear decoder (blue lines); forelimb displacement as predicted by an LSTM nonlinear decoder (orange lines). Shaded blue regions represent periods of active sleep.

(B) Mean (+SEM) overall performance of the linear (blue bars) and LSTM (orange bars) decoders for P8 and P12 rats, as measured by $r^2$ (the proportion of variance in forelimb displacement explained by the decoder). Brackets denote a significant interaction between age and decoder ($p < 0.05$). Asterisk denotes significantly better performance of the linear decoder at P8 compared with at P12 ($p < 0.05$).

(C) Left: same as in (B) but for periods of forelimb movement. Right: same as in (B) but for nonmovement periods; brackets denote that both decoders performed significantly better at P8 than at P12 ($p < 0.05$).

In contrast, because continuous activity at P12 occludes the temporal and spatial relations between M1 activity and forelimb movements, we predicted that a linear decoder would no longer predict forelimb movements at that age. Indeed, by comparing actual with predicted limb displacement using a linear decoder (Figure 1A, blue lines), we confirmed both predictions: a linear decoder is sufficient to predict forelimb movements at P8 but not at P12.

The observed loss of decoding accuracy at P12 must be reconciled with the fact that M1 activity is highly correlated with limb kinematics in adults.[23,24] Accordingly, we hypothesized that M1 activity continues to represent movement kinematics after the emergence of continuous activity but that this representation can no longer be captured by a linear decoder. We tested this hypothesis using a nonlinear decoder (long short-term memory [LSTM] decoder[26]) and found that it effectively decodes forelimb movements at P12, as well as at P8 (Figure 1A, orange lines).

To quantify decoder performance, we computed the proportion of variance explained ($r^2$) between the actual and predicted forelimb displacement (Figure 1B). This metric represents the amount of temporal and spatial kinematic information available in M1 during a forelimb movement. An ANOVA revealed a significant age × decoder interaction ($F(1, 14) = 19.15$, $p < 0.001$, adjusted [adj.] $\eta_p^2 = 0.548$; Figure 1B), indicating that the linear decoder performed significantly better at P8 than at P12 ($F(1, 14) = 10.08$, $p = 0.007$, adj. $\eta_p^2 = 0.376$) and that the nonlinear decoder significantly outperformed the linear decoder at both ages. Thus, M1 activity indeed preserves kinematic information after continuous activity emerges at P12, but this information is only observable using a nonlinear decoder.

## State-dependent encoding of movement-related information

Next, we confirmed that decoder performance was driven by M1 activity specifically during periods of limb movement. Decoding was performed across the entire recording period, after which we segmented the data into movement and nonmovement periods as described previously.[6] Successful decoder performance was attributable to movement periods ($F(1, 14) = 7.35$, $p = 0.017$, adj. $\eta_p^2 = 0.297$; Figure 1C, left) and not to nonmovement periods ($F(1, 14) = 0.06$, $p = 0.817$, adj. $\eta_p^2 < 0.001$; Figure 1C, right). Although there was a main effect of age on decoder performance during nonmovement periods ($F(1, 14) = 9.15$, $p = 0.009$, adj. $\eta_p^2 = 0.352$), this effect was negligible compared to the performance of the decoder during movement periods.

Because the previous analysis of movement periods did not distinguish between twitches and wake movements, we next assessed decoder performance separately for each type of movement (Figures 2A, S1A, and S1B). For this analysis, we used a 2-s window centered on each twitch and wake movement. For twitches, a main effect of age indicated that decoder performance was significantly better at P8 than at P12, regardless of

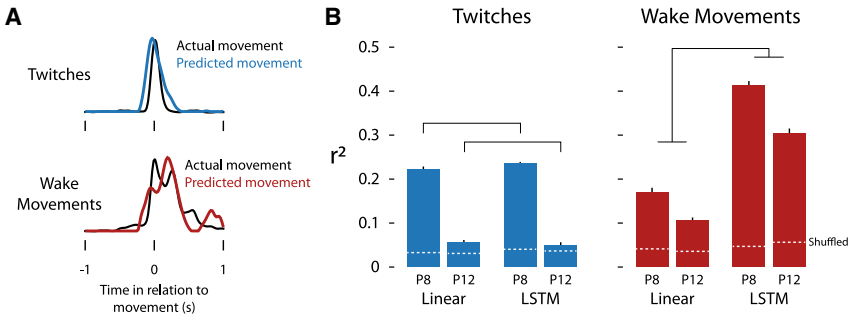



**Figure 2. Decoder performance for twitches and wake movements varies by age**

(A) Forelimb displacement for actual movements (black lines) compared with a predicted twitch (blue line, top) and a predicted wake movement (red line, bottom). A single representative twitch and wake movement is shown.

(B) Mean (+SEM) decoder performance ($r^2$) for twitches (left; blue bars) and wake movements (right; red bars). For each type of movement, the linear decoder is compared with the LSTM decoder across P8 and P12 rats. For twitches, brackets denote that both decoders performed significantly better at P8 than at P12 ($p < 0.05$). For wake movements, brackets denote that the LSTM decoder performed significantly better than the linear decoder at both ages ($p < 0.05$). Horizontal dashed white lines indicate chance performance using a shuffling procedure.

which decoder was used ($F(1, 3,808) = 534.25$, $p < 0.001$, adj. $\eta_p^2 = 0.123$; Figure 2B, left). Unexpectedly, for twitches, the nonlinear decoder's performance did not improve on the linear decoder's performance at either age ($F(1, 3,808) = 1.67$, $p = 0.197$, adj. $\eta_p^2 < 0.001$). This finding suggests that M1 activity no longer represents twitch kinematics after continuous activity emerges (even though reafference from twitches continues to trigger M1 activity at P12 and older ages).[27]

In contrast to twitches, for wake movements, the nonlinear decoder performed significantly better than the linear decoder at P8 and P12 ($F(1, 1,340) = 340.91$, $p < 0.001$, adj. $\eta_p^2 = 0.202$; Figure 2B, right). This finding is particularly surprising at P8 because reafference from wake movements is blunted relative to twitches at this age.[20,21] Thus, wake movements appear to be represented in a nonlinear fashion as early as P8, and this representation persists despite the transition to continuous activity at P12. For both decoders, there was a small but statistically significant decrease in performance at P12 compared with at P8 ($F(1, 1,340) = 18.24$, $p < 0.001$, adj. $\eta_p^2 = 0.012$), suggesting that the emergence of continuous activity modestly interferes with M1's representation of wake movements.

It is possible that by decoding across the entire recording period, the decoder model was biased toward predicting wake movements at the expense of twitches, leading to relatively poor twitch predictions at P12. However, we found that decoding sleep and wake periods separately did not improve decoder performance for twitches at either age (Figure S1C).

Altogether, these results indicate that although M1's representation of twitch kinematics diminishes between P8 and P12, its representation of wake-movement kinematics is robust at both ages.

## Continuous activity increases the uniqueness of M1 information

The presence of decorrelated, continuous M1 activity at P12 represents a transition from a "dense" (i.e., redundant) encoding of information to a more energy- and information-efficient "sparse" code.[28] Accordingly, we predicted that the random removal of neural input would degrade decoder performance more at P12 than at P8.

To test this hypothesis, two "lesion" experiments were performed. In the first "unit lesion" experiment, a varying percentage of M1 units (up to 90%) recorded from a given pup was randomly discarded from the decoding process (Figure 3A, top). The resulting decoder performance was then compared with the decoder performance of the "intact" model (i.e., 0% of units removed). (Only the nonlinear decoder was used for these experiments.) As predicted, although unit lesions degraded decoder performance at both ages, the loss was significantly greater at P12 than at P8 ($F(1.50, 20.94) = 18.60$, $p < 0.001$, adj. $\eta_p^2 = 0.539$; Figure 3A, bottom). For example, lesioning 50% of units led to a decoder performance of 70.4% ± 3.9% at P8, compared with 41.9% ± 4.4% at P12. Differences in decoder performance were statistically significant across the range of unit lesions (all $p \leq 0.005$).

The second "spike lesion" experiment was performed in a similar fashion. Here, a varying percentage of spikes (i.e., action potentials) across all M1 units (up to 90%) was randomly discarded from the decoding process (Figure 3B, top). As predicted, spike lesions resulted in significantly greater declines in decoder performance at P12 than at P8 ($F(5.05, 70.63) = 8.41$, $p < 0.001$, adj. $\eta_p^2 = 0.330$; Figure 3B, bottom). For example, lesioning 50% of spikes led to a decoder performance of 79.5% ± 1.8% at P8, compared with 51.0% ± 3.2% at P12. Again, differences in decoder performance were statistically significant across the range of spikes removed (all $p \leq 0.03$).

That both lesions decreased decoder performance more at P12 than at P8 implies that each M1 unit's activity contains more unique information about limb kinematics at the older age. To test this implication, we measured the mutual information between a target M1 unit and a randomly selected subset of units[29]; the size of the subset varied from 1 additional unit to all available units for a given pup. Here, mutual information refers to the predictability of the target unit's response (to a movement) when the responses of the subset are known. By definition, mutual information only increases when the subset contains nonredundant (i.e., unique) information about the target unit.[30] We found that adding units increased mutual information significantly more quickly at P12 (38.7% ± 1% units to reach 0.5 bits) than at P8 (50.9% ± 1.2% units to reach 0.5 bits; $t(436.01) = 7.64$; $p \leq 0.001$, adj. $\eta^2 = 0.116$; Figure 3C; note that 1 bit is the theoretical maximum value). Together, these results support the notion that the onset of continuous activity at P12 increases the uniqueness of information available in M1.

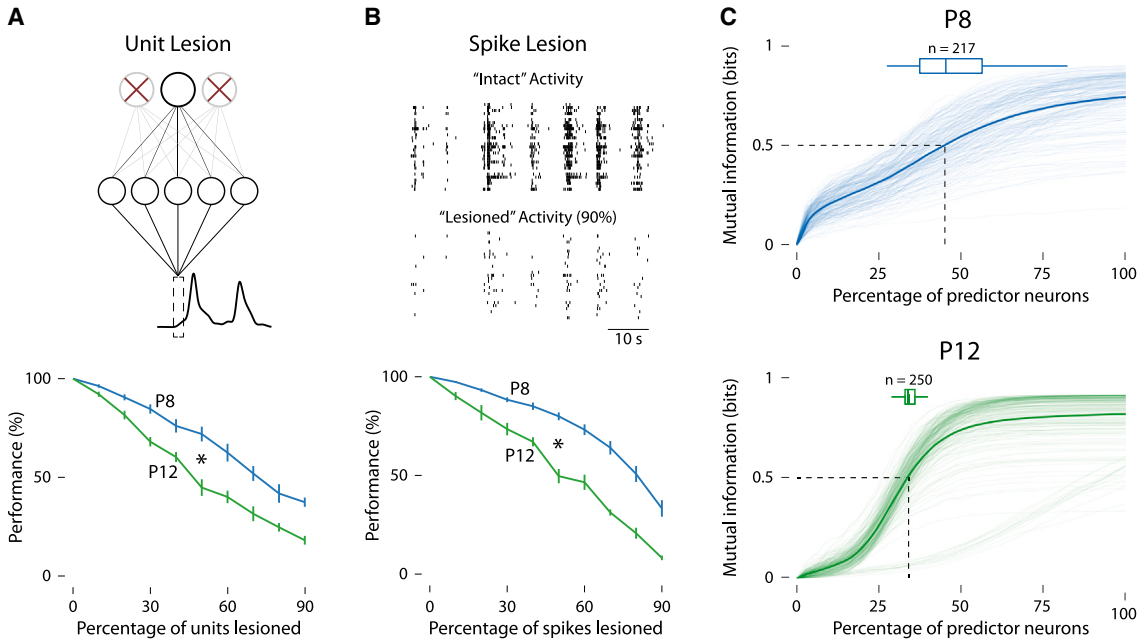



**Figure 3. M1 activity contains more unique information at P12 than at P8**

(A) Top: representation of the "unit lesion" experiment. The top row of nodes represents the individual M1 units input to the LSTM decoder. Red Xs denote units that were "lesioned" (i.e., removed from the decoding process). The second row of nodes represents the hidden layer of the LSTM decoder. The black trace at the bottom represents forelimb displacement as predicted by the LSTM decoder. Bottom: mean (±SEM) decoder performance (measured as the percentage change in the "lesioned" $r^2$ relative to the "intact" $r^2$) for P8 (blue line) and P12 (green line) rats as a function of the percentage of units lesioned. Asterisk denotes significantly worse decoder performance at P12 than at P8 across the range of units lesioned ($p < 0.05$).

(B) Top: representation of the "spike lesion" experiment. "Intact" activity represents baseline M1 activity, where each row represents an individual unit and each vertical tick denotes an action potential. "Lesioned" activity represents M1 activity after 90% of individual spikes were removed from the decoding process. Bottom: same as in (A) but for the spike lesion experiment. Asterisk denotes significantly worse decoder performance at P12 than at P8 across the range of spikes lesioned ($p < 0.05$).

(C) For P8 (top, blue) and P12 (bottom, green) rats, mutual information (in bits) is shown as a function of increasing size of the predictor subset (expressed as a percentage of the maximum size). Solid lines represent the mean increase in mutual information in the target unit; translucent lines represent individual target units (P8: n = 217 units; P12: n = 250 units). Dashed lines show the mean percentage of predictor units required to achieve 0.5 bits (half the theoretical maximum); horizontal boxplots show the percentage of predictor units required to achieve 0.5 bits for individual target units.

## M1's encoding scheme is more "generic" at P8 than at P12

The finding that M1's representation of movement at P8, but not at P12, is so redundant as to be robust to the removal of neural activity raises an intriguing hypothesis: namely, that discontinuous activity is not only redundant but also generic to the point of being interchangeable between individual pups. Such a property could reflect a gross encoding scheme during early development when somatotopic relations among M1 units and forelimb muscles are still being established. Accordingly, we predicted that after "transplanting" the encoding scheme of one pup into that of another, decoder performance would remain high at P8 but would be degraded at P12 (Figure 4A).

To test this prediction, M1 activity for a "recipient" pup was decoded using the model weights (inferred encoding scheme) transplanted from a same-age "donor" pup. M1 units were aligned between the donor and recipient pups according to firing rate (i.e., sorted from highest to lowest firing rates). For both P8 and P12 rats, 29 such donor-recipient pairs were generated. (Again, only the nonlinear decoder was used.) As predicted, decoder performance was significantly better across P8 pairings than across P12 pairings ($t(56) = 11.71$, $p < 0.001$, adj. $\eta^2 = 0.705$; Figures 4B and 4C). To ensure that this finding was not simply due to our method of aligning firing rates across pups, we performed 30 additional tests using random unit-unit pairings; this randomization had no substantive effect on the age-related difference in decoder performance after transplantation (Figure 4D). Thus, continuous activity is associated with increased individualization of M1 encoding schemes at P12.

## DISCUSSION

We previously reported that M1 units encode movement kinematics—especially twitch amplitude—at P8 but that this encoding disappears with the onset of continuous activity at P12.[6] That finding raised the possibility that M1's somatosensory representations are "reset" at P12 by the sudden change in cortical dynamics. An alternative possibility was that the somatosensory representation persists at P12 but is obscured by the continuous activity. We considered the second possibility more likely than the first; accordingly, we predicted that a nonlinear decoder

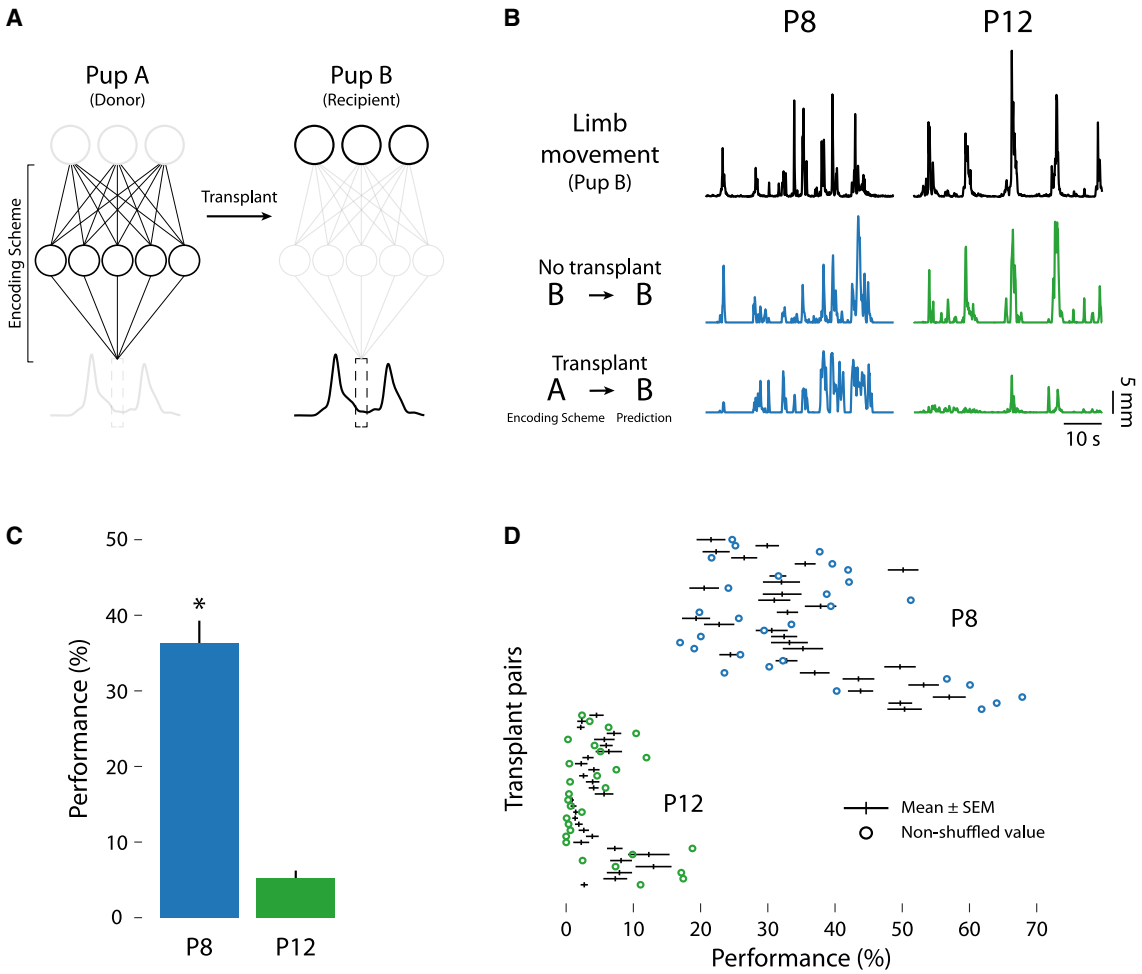



**Figure 4. Transplanting M1's encoding scheme between pups leads to successful decoding at P8 but not at P12**

(A) Representation of the "model transplant" experiment. Pup A ("donor"; left) donates its encoding scheme to pup B ("recipient"; right). Thus, pup B's M1 unit activity (top row) is used to predict pup B's forelimb displacement (black trace at bottom) using pup A's encoding scheme.

(B) Representative traces of actual limb movement (top) compared with limb movements predicted by pup B's (original) decoder model (middle) and limb movement predicted by pup B's (transplanted) decoder model (bottom) for P8 (left, blue lines) and P12 (right, green lines) rats.

(C) Mean (+SEM) decoder performance (measured as the ratio of the "transplant" $r^2$ to the original $r^2$, expressed as a percentage) for P8 (blue) and P12 (green) rats. Asterisk denotes significantly better decoder performance after transplantation at P8 than at P12 ($p < 0.05$).

(D) Decoder performance (now on the x axis) is shown for each donor-recipient pair across 30 random shuffles. Black lines indicate the mean (±SEM) decoder performance of the 30 random shuffles for each donor-recipient pair. Blue and green circles indicate decoder performance of the unshuffled models at P8 and P12, respectively.

would reveal that M1 continues to encode movement-related information at P12.

Our prediction was confirmed at P12 for wake movements but not for twitches. The latter result surprised us because twitch-related reafference continues to trigger M1 activity at P12 and beyond.[27] Moreover, wake movements were better reconstructed by the nonlinear decoder at P8, indicating that a significant component of M1's representation of wake movements is nonlinear even before the onset of continuous activity. This last finding also surprised us because, at this age, the brainstem selectively dampens (though does not eliminate) wake-related reafference in M1[20,21,31]; also, our previous analysis using a linear model led us to conclude that M1 units are not sensitive to wake-related kinematics at either age.[6] Thus, the present finding that, at P8, M1 encodes twitches and wake movements through linear and nonlinear means, respectively, adds another dimension to our understanding of M1's state-dependent sensory representation in early development.

It is not clear why nonlinear decoder performance was so much better for wake movements than for twitches. The explanation may reside in the fact that wake movements, which involve multiple and prolonged muscle activations, provide richer spatiotemporal information for neural decoding than twitches, which are discrete and brief. Conversely, the unique spatiotemporal features of twitches make them better suited than wake movements for developing and refining sensorimotor

circuits.[32] Moreover, that twitches were not as reliably decoded as wake movements at P12 does not suggest an end to their functional significance. For example, in P20 rats, twitch-related neural activity is implicated in the developmental emergence of a cerebellar-dependent internal model of movement.[27] To what extent twitching contributes to other forms of plasticity in adults remains an open question.[33–35]

### The onset of continuous activity is associated with more unique M1 information

In adults, continuous activity is associated with complex functions such as sparse coding[28,36,37] and predictive coding.[36,38,39] Continuous activity is also thought to enhance reafference by providing contextual information related to behavior, cognition, and cortical dynamics.[40–42] Accordingly, one might expect the emergence of continuous activity at P12 to immediately sharpen reafference in M1. However, this was not the case (see Figure 2B), suggesting that continuous activity per se does not enhance reafference or lead to better neural decoding outcomes.

Although continuous activity did not immediately enhance reafference at P12, it did correspond with an increase in the uniqueness of information in M1. This increase in unique information could be due, in part, to the developmental narrowing of M1 receptive fields. At P8, M1 units respond across a range of movement amplitudes, suggesting that M1 receptive fields are broadly tuned to multiple forelimb muscles.[6] This broad tuning is no longer apparent at P12, suggesting a progression toward the narrow receptive fields displayed by M1 units in adulthood.[43–45] Similar receptive-field narrowing has been demonstrated in other cortical[46,47] and subcortical[48,49] areas and is an activity-dependent process.[50]

If the increased continuous activity at P12 is accurately characterized as noise, then that noise could function to promote a process akin to "regularization."[51] Regularization entails the introduction of noise to a system to prevent overfitting (i.e., forming inappropriately strong relationships too early in the learning process). Regularization through noise is a conventional technique in the creation of robust computational networks.[43,52,53] Accordingly, we propose that regularization contributes to receptive-field narrowing by ensuring that connections between M1 units and muscle fibers are only strengthened when they fire together with high-temporal precision. Connections with less-precise temporal relations (i.e., at the edges of the receptive field) are weakened and pruned.

### M1's encoding scheme is more individualized at P12

We found that when the encoding scheme of one P8 rat was "transplanted" to another P8 rat, decoder performance was significantly better than after a similar "transplant" between P12 rats. In other words, M1's encoding schemes at P12 are more complex and more individualized than the schemes at P8.

Developing animals face the problem of matching cortical connections to a moving target: their rapidly growing bodies. Discontinuous activity may help to solve this problem before P12, as nearly all M1 activity occurs in response to movement-related reafference or external stimulation.[6,8,54] This pattern of activity maximizes the correlation between behavior and neural activity. Similarly, before P12, a 10- to 20-Hz corticothalamic rhythm—known as a spindle burst—amplifies reafference and promotes the development of sensorimotor pathways.[2,55–58] This strong association between limb movements and M1 activity is thought to strengthen early-developing M1 connections that serve as the foundation for later-emerging motor control.[6,21,54]

### Limitations of the study

One limitation of this study is that all analyses were performed on pre-recorded (i.e., static) spike trains, and thus we were not able to assess how M1's encoding scheme dynamically responds to perturbations of individual units or neural ensembles. Another limitation is the exclusive focus on two ages surrounding the transition to continuous activity in one cortical area. Future work should expand the developmental approach to neural decoding introduced here to older ages and additional cortical areas.

### Conclusion

In adults, M1 contributes in many complex ways to motor control and motor learning.[59–61] Also, M1 integrates information arising from the other senses,[62,63] neuromodulatory systems,[64,65] and behavior.[40–42] Although little is known about the development of these higher-order functions of M1, it is now clear that these functions—like M1's most basic motor capacities—rest upon an early-developing sensory foundation. But even this early "sensory phase" of M1 development is protracted and complex: it begins during the discontinuous period of M1 activity as gross somatotopic relations are formed and, as shown here, is transformed through the transition to continuous activity. Thus, the present findings add new dimensions to our growing understanding of M1's sensory development before it assumes its more familiar role in motor control.

## STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS
- METHOD DETAILS
- QUANTIFICATION AND STATISTICAL ANALYSIS
  - Data preparation
  - Data scaling
  - Linear model
  - Nonlinear model
  - Decoder performance
  - Model lesion experiments
  - Mutual information
  - Model transplant experiment
  - Statistical analyses

## SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.celrep.2023.113119.

## ACKNOWLEDGMENTS

Preparation of this article was made possible by a grant from the National Institute of Child Health and Human Development (R37-HD081168) to M.S.B.

## AUTHOR CONTRIBUTIONS

Conceptualization, R.M.G. and M.S.B.; methodology, R.M.G.; software, R.M.G.; formal analysis, R.M.G.; investigation, R.M.G.; data curation, R.M.G.; writing – original draft, R.M.G.; writing – review & editing, R.M.G., G.S., and M.S.B.; visualization, R.M.G. and M.S.B.; funding acquisition, G.S. and M.S.B.; resources, G.S. and M.S.B.; supervision, M.S.B.

## DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: February 1, 2023
Revised: June 8, 2023
Accepted: August 24, 2023

## REFERENCES

1. Vanhatalo, S., and Kaila, K. (2006). Development of neonatal EEG activity: from phenomenology to physiology. Semin. Fetal Neonatal Med. *11*, 471–478.
2. Colonnese, M.T., and Phillips, M.A. (2018). Thalamocortical function in developing sensory circuits. Curr. Opin. Neurobiol. *52*, 72–79.
3. Vanhatalo, S., Palva, J.M., Andersson, S., Rivera, C., Voipio, J., and Kaila, K. (2005). Slow endogenous activity transients and developmental expression of K+–Cl− cotransporter 2 in the immature human cortex. Eur. J. Neurosci. *22*, 2799–2804.
4. Golshani, P., Gonçalves, J.T., Khoshkhoo, S., Mostany, R., Smirnakis, S., and Portera-Cailliau, C. (2009). Internally mediated developmental desynchronization of neocortical network activity. J. Neurosci. *29*, 10890–10899.
5. Rochefort, N.L., Garaschuk, O., Milos, R.-I., Narushima, M., Marandi, N., Pichler, B., Kovalchuk, Y., and Konnerth, A. (2009). Sparsification of neuronal activity in the visual cortex at eye-opening. Proc. Natl. Acad. Sci. USA *106*, 15049–15054.
6. Glanz, R.M., Dooley, J.C., Sokoloff, G., and Blumberg, M.S. (2021). Sensory coding of limb kinematics in motor cortex across a key developmental transition. J. Neurosci. *41*, 6905–6918.
7. Riyahi, P., Phillips, M.A., and Colonnese, M.T. (2021). Input-independent homeostasis of developing thalamocortical activity. eNeuro *8*. ENEURO.0184-21.2021.
8. Gómez, L.G., Dooley, J.C., and Blumberg, M.S. (2023). Activity in developing prefrontal cortex is shaped by sleep and sensory experience. Elife.
9. Virtanen, M.A., Uvarov, P., Mavrovic, M., Poncer, J.C., and Kaila, K. (2021). The multifaceted roles of KCC2 in cortical development. Trends Neurosci. *44*, 378–392.
10. Khazipov, R., Khalilov, I., Tyzio, R., Morozova, E., Ben-Ari, Y., and Holmes, G.L. (2004). Developmental changes in GABAergic actions and seizure susceptibility in the rat hippocampus. Eur. J. Neurosci. *19*, 590–600.
11. Butt, S.J., Stacey, J.A., Teramoto, Y., and Vagnoni, C. (2017). A role for GABAergic interneuron diversity in circuit development and plasticity of the neonatal cerebral cortex. Curr. Opin. Neurobiol. *43*, 149–155.
12. Alcántara, S., Ferrer, I., and Soriano, E. (1993). Postnatal development of parvalbumin and calbindin D28K immunoreactivities in the cerebral cortex of the rat. Anat. Embryol. *188*, 63–73.
13. Murata, Y., and Colonnese, M.T. (2019). Thalamic inhibitory circuits and network activity development. Brain Res. *1706*, 13–23.
14. Dard, R.F., Leprince, E., Denis, J., Rao Balappa, S., Suchkov, D., Boyce, R., Lopez, C., Giorgi-Kurz, M., Szwagier, T., Dumont, T., et al. (2022). The rapid developmental rise of somatic inhibition disengages hippocampal dynamics from self-motion. Elife *11*, e78116.
15. Downes, N., and Mullins, P. (2014). The development of myelin in the brain of the juvenile rat. Toxicol. Pathol. *42*, 913–922.
16. Jouvet-Mounier, D., Astic, L., and Lacote, D. (1970). Ontogenesis of the states of sleep in rat, cat, and guinea pig during the first postnatal month. Dev. Psychobiol. *2*, 216–239.
17. Gramsbergen, A. (1976). The development of the EEG in the rat. Dev. Psychobiol. *9*, 501–515.
18. Seelke, A.M.H., and Blumberg, M.S. (2008). The microstructure of active and quiet sleep as cortical delta activity emerges in infant rats. Sleep *31*, 691–699.
19. Del Rio-Bermudez, C., Kim, J., Sokoloff, G., and Blumberg, M.S. (2017). Theta oscillations during active sleep synchronize the developing rubrohippocampal sensorimotor network. Curr. Biol. *27*, 1413–1424.e4.
20. Tiriac, A., Del Rio-Bermudez, C., and Blumberg, M.S. (2014). Self-generated movements with "unexpected" sensory consequences. Curr. Biol. *24*, 2136–2141.
21. Dooley, J.C., and Blumberg, M.S. (2018). Developmental 'awakening' of primary motor cortex to the sensory consequences of movement. Elife *7*, e41841.
22. Singleton, A.C., Brown, A.R., and Teskey, G.C. (2021). Development and plasticity of complex movement representations. J. Neurophysiol. *125*, 628–637.
23. Ashe, J., and Georgopoulos, A.P. (1994). Movement parameters and neural activity in motor cortex and area 5. Cereb. Cortex *4*, 590–600.
24. Georgopoulos, A.P., Kettner, R.E., and Schwartz, A.B. (1988). Primate motor cortex and free arm movements to visual targets in three-dimensional space. II. Coding of the direction of movement by a neuronal population. J. Neurosci. *8*, 2928–2937.
25. Glaser, J.I., Benjamin, A.S., Chowdhury, R.H., Perich, M.G., Miller, L.E., and Kording, K.P. (2020). Machine learning for neural decoding. eNeuro *7*. ENEURO.0506-19.2020.
26. Hochreiter, S., and Schmidhuber, J. (1997). Long short-term memory. Neural Comput. *9*, 1735–1780.
27. Dooley, J.C., Sokoloff, G., and Blumberg, M.S. (2021). Movements during sleep reveal the developmental emergence of a cerebellar-dependent internal model in motor thalamus. Curr. Biol. *31*, 5501–5511.e5.
28. Olshausen, B.A., and Field, D.J. (2004). Sparse coding of sensory inputs. Curr. Opin. Neurobiol. *14*, 481–487.
29. Shannon, C.E. (1948). A mathematical theory of communication. Bell Syst. Tech. J. *27*, 379–423.
30. Bettencourt, L.M.A. (2009). The rules of information aggregation and emergence of collective intelligent behavior. Top. Cogn. Sci. *1*, 598–620.
31. Tiriac, A., and Blumberg, M.S. (2016). Gating of reafference in the external cuneate nucleus during self-generated movements in wake but not sleep. Elife *5*, e18749.
32. Blumberg, M.S., Dooley, J.C., and Tiriac, A. (2022). Sleep, plasticity, and sensory neurodevelopment. Neuron *110*, 3230–3242.
33. Herman, J.H., and Roffwarg, H.P. (1983). Modifying oculomotor activity in awake subjects increases the amplitude of eye movements during REM sleep. Science *220*, 1074–1076.
34. De Gennaro, L., Ferrara, M., Urbani, L., Bertini, M., and Neuroscienze, S.D. (2000). A complementary relationship between wake and REM sleep in the auditory system: a pre-sleep increase of middle-ear muscle activity

(MEMA) causes a decrease of MEMA during sleep. Exp. Brain Res. *130*, 105–112.

35. Eckert, M.J., McNaughton, B.L., and Tatsuno, M. (2020). Neural ensemble reactivation in rapid eye movement and slow-wave sleep coordinate with muscle activity to promote rapid motor skill learning. Philos. Trans. R. Soc. Lond. B Biol. Sci. *375*, 20190655.
36. Atick, J.J. (1992). Could information theory provide an ecological theory of sensory processing? Netw. Comput. Neural Syst. *3*, 213–251.
37. Stringer, C., Pachitariu, M., Steinmetz, N., Reddy, C.B., Carandini, M., and Harris, K.D. (2019). Spontaneous behaviors drive multidimensional, brain-wide activity. Science *364*, 255.
38. van Hateren, J.H. (1992). A theory of maximizing sensory information. Biol. Cybern. *68*, 23–29.
39. Keller, G.B., and Mrsic-Flogel, T.D. (2018). Predictive processing: a canonical cortical computation. Neuron *100*, 424–435.
40. Arieli, A., Sterkin, A., Grinvald, A., and Aertsen, A. (1996). Dynamics of ongoing activity: explanation of the large variability in evoked cortical responses. Science *273*, 1868–1871.
41. Goris, R.L.T., Movshon, J.A., and Simoncelli, E.P. (2014). Partitioning neuronal variability. Nat. Neurosci. *17*, 858–865.
42. Lin, I.-C., Okun, M., Carandini, M., and Harris, K.D. (2015). The nature of shared cortical variability. Neuron *87*, 644–656.
43. Stefanis, C., and Jasper, H. (1964). Intracellular microelectrode studies of antidromic responses in cortical pyramidal tract neurons. J. Neurophysiol. *27*, 828–854.
44. Nudo, R.J., Jenkins, W.M., Merzenich, M.M., Prejean, T., and Grenda, R. (1992). Neurophysiological correlates of hand preference in primary motor cortex of adult squirrel monkeys. J. Neurosci. *12*, 2918–2947.
45. Georgopoulos, A.P., and Stefanis, C.N. (2007). Local shaping of function in the motor cortex: Motor contrast, directional tuning. Brain Res. Rev. *55*, 383–389.
46. Hubel, D.H., and Wiesel, T.N. (1963). Receptive fields of cells in striate cortex of very young, visually inexperienced kittens. J. Neurophysiol. *26*, 994–1002.
47. DeAngelis, G.C., Ohzawa, I., and Freeman, R.D. (1993). Spatiotemporal organization of simple-cell receptive fields in the cat's striate cortex. I. General characteristics and postnatal development. J. Neurophysiol. *69*, 1091–1117.
48. Chen, C., and Regehr, W.G. (2000). Developmental remodeling of the retinogeniculate synapse. Neuron *28*, 955–966.
49. Tschetter, W.W., Govindaiah, G., Etherington, I.M., Guido, W., and Niell, C.M. (2018). Refinement of spatial receptive fields in the developing mouse lateral geniculate nucleus is coordinated with excitatory and inhibitory remodeling. J. Neurosci. *38*, 4531–4542.
50. Chakrabarty, S., and Martin, J.H. (2005). Motor but not sensory representation in motor cortex depends on postsynaptic activity during development and in maturity. J. Neurophysiol. *94*, 3192–3198.
51. Camuto, A., Willetts, M., Simsekli, U., Roberts, S.J., and Holmes, C.C. (2020). Explicit regularisation in gaussian noise injections. Adv. Neural Inf. Process. Syst. *33*, 16603–16614.
52. Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., and Salakhutdinov, R. (2014). Dropout: a simple way to prevent neural networks from overfitting. J. Mach. Learn. Res. *15*, 1929–1958.
53. Poole, B., Sohl-Dickstein, J., and Ganguli, S. (2014). Analyzing noise in autoencoders and deep networks. Preprint at arXiv. https://doi.org/10.48550/arXiv.1406.1831.
54. Gómez, L.J., Dooley, J.C., Sokoloff, G., and Blumberg, M.S. (2021). Parallel and serial sensory processing in developing primary somatosensory and motor cortex. J. Neurosci. *41*, 3418–3431.
55. Khazipov, R., Sirota, A., Leinekugel, X., Holmes, G.L., Ben-Ari, Y., and Buzsáki, G. (2004). Early motor activity drives spindle bursts in the developing somatosensory cortex. Nature *432*, 758–761.
56. Tolner, E.A., Sheikh, A., Yukin, A.Y., Kaila, K., and Kanold, P.O. (2012). Subplate neurons promote spindle bursts and thalamocortical patterning in the neonatal rat somatosensory cortex. J. Neurosci. *32*, 692–702.
57. Yang, J.-W., An, S., Sun, J.-J., Reyes-Puerta, V., Kindler, J., Berger, T., Kilb, W., and Luhmann, H.J. (2013). Thalamic network oscillations synchronize ontogenetic columns in the newborn rat barrel cortex. Cereb. Cortex *23*, 1299–1316.
58. Dooley, J.C., Glanz, R.M., Sokoloff, G., and Blumberg, M.S. (2020). Self-generated whisker movements drive state-dependent sensory input to developing barrel cortex. Curr. Biol. *30*, 2404–2410.e4.
59. Sanes, J.N., and Donoghue, J.P. (2000). Plasticity and primary motor cortex. Annu. Rev. Neurosci. *23*, 393–415.
60. Kawai, R., Markman, T., Poddar, R., Ko, R., Fantana, A.L., Dhawale, A.K., Kampff, A.R., and Ölveczky, B.P. (2015). Motor cortex is required for learning but not for executing a motor skill. Neuron *86*, 800–812.
61. Peters, A.J., Liu, H., and Komiyama, T. (2017). Learning in the rodent motor cortex. Annu. Rev. Neurosci. *40*, 77–97.
62. Ibrahim, L.A., Huang, S., Fernandez-Otero, M., Sherer, M., Qiu, Y., Vemuri, S., Xu, Q., Machold, R., Pouchelon, G., Rudy, B., and Fishell, G. (2021). Bottom-up inputs are required for establishment of top-down connectivity onto cortical layer 1 neurogliaform cells. Neuron *109*, 3473–3485.e5.
63. Schuman, B., Dellal, S., Prönneke, A., Machold, R., and Rudy, B. (2021). Neocortical layer 1: an elegant solution to top-down and bottom-up integration. Annu. Rev. Neurosci. *44*, 221–252.
64. Mechawar, N., Cozzari, C., and Descarries, L. (2000). Cholinergic innervation in adult rat cerebral cortex: a quantitative immunocytochemical description. J. Comp. Neurol. *428*, 305–318.
65. Allaway, K.C., Muñoz, W., Tremblay, R., Sherer, M., Herron, J., Rudy, B., Machold, R., and Fishell, G. (2020). Cellular birthdate predicts laminar and regional cholinergic projection topography in the forebrain. Elife *9*, e63249.
66. 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. Nat. Neurosci. *21*, 1281–1289.
67. Nath, T., Mathis, A., Chen, A.C., Patel, A., Bethge, M., and Mathis, M.W. (2019). Using DeepLabCut for 3D markerless pose estimation across species and behaviors. Nat. Protoc. *14*, 2152–2176.
68. Varma, S., and Simon, R. (2006). Bias in error estimation when using cross-validation for model selection. BMC Bioinf. *7*, 91.
69. Abadi, M., Barham, P., Chen, J., Chen, Z., Davis, A., Dean, J., Devin, M., Ghemawat, S., Irving, G., and Isard, M. (2016). TensorFlow: a system for large-scale machine learning. In 12th USENIX Symposium on Operating Systems Design and Implementation (OSDI 16).
70. Rosenblatt, F. (1958). The perceptron: a probabilistic model for information storage and organization in the brain. Psychol. Rev. *65*, 386–408.
71. Chollet, F. (2018). Keras: the python Deep Learning Library (Astrophysics Source Code Library).
72. Kingma, D.P., and Ba, J. (2014). Adam: A method for stochastic optimization. Preprint at arXiv. https://doi.org/10.48550/arXiv.1412.6980.
73. Wang, Y., Truccolo, W., and Borton, D.A. (2018). Decoding hindlimb kinematics from primate motor cortex using long short-term memory recurrent neural networks. In 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC).
74. Naufel, S., Glaser, J.I., Kording, K.P., Perreault, E.J., and Miller, L.E. (2019). A muscle-activity-dependent gain between motor cortex and EMG. J. Neurophysiol. *121*, 61–73.
75. Cortes, C., Mohri, M., and Rostamizadeh, A. (2012). L2 regularization for learning kernels. Preprint at arXiv. https://doi.org/10.48550/arXiv.1205.2653.
76. Hyndman, R.J., and Koehler, A.B. (2006). Another look at measures of forecast accuracy. Int. J. Forecast. *22*, 679–688.

77. Harris, C.R., Millman, K.J., Van Der Walt, S.J., Gommers, R., Virtanen, P., Cournapeau, D., Wieser, E., Taylor, J., Berg, S., Smith, N.J., et al. (2020). Array programming with NumPy. Nature *585*, 357–362.

78. Schneidman, E., Still, S., Berry, M.J., and Bialek, W. (2003). Network information and connected correlations. Phys. Rev. Lett. *91*, 238701.

79. Bettencourt, L.M.A., Gintautas, V., and Ham, M.I. (2008). Identification of functional information subgraphs in complex networks. Phys. Rev. Lett. *100*, 238701.

80. Timme, N., Alford, W., Flecker, B., and Beggs, J.M. (2014). Synergy, redundancy, and multivariate information measures: an experimentalist's perspective. J. Comput. Neurosci. *36*, 119–140.

81. Timme, N.M., and Lapish, C. (2018). A tutorial for information theory in neuroscience. eNeuro *5*. ENEURO.0052-18.2018.

82. Mordkoff, J.T. (2019). A simple method for removing bias from a popular measure of standardized effect size: Adjusted partial eta squared. Adv. Methods Pract. Psychol. Sci. *2*, 228–232.

## STAR★METHODS

### KEY RESOURCES TABLE

| REAGENT or RESOURCE | SOURCE | IDENTIFIER |
|---|---|---|
| Experimental models: Organisms/strains | | |
| Sprague-Dawley Norway Rats | Envigo | RRID: RGD_10401918 |
| Software and algorithms | | |
| MATLAB, version 2020b | Mathworks | RRID: SCR_001622 |
| Python Programming Language, version 3.8 | Python Software Foundation | RRID: SCR_008394 |
| Tensorflow, version 2.9 | Google | RRID: SCR_016345 |
| Adobe Illustrator Creative Cloud 2022 | Adobe | RRID: SCR_010279 |
| SPSS 28 | IBM | RRID:SCR_019096 |
| Other | | |
| Time-stamped sleep and movement data | This paper | https://github.com/rglanz/Glanz_et_al_2023 https://zenodo.org/record/8238644 |

### RESOURCE AVAILABILITY

#### Lead contact

Further information and requests for resources should be directed to, and will be fulfilled by, the lead contact, Dr. Mark Blumberg (mark-blumberg@uiowa.edu).

#### Materials availability

This study did not generate new unique reagents.

#### Data and code availability

Timestamps of action potentials, twitches, wake movements, and sleep-wake transitions, as well as the position of the forelimb in Cartesian coordinates are available for download on Github (https://github.com/rglanz/Glanz_et_al_2023) and Zenodo (https://zenodo.org/record/8238644, https://doi.org/10.5281/zenodo.8238644). All custom software is available upon request. (Also see https://github.com/KordingLab/Neural_Decoding for neural decoding software, and https://github.com/nmtimme/Neuroscience-Information-Theory-Toolbox for information theory software.) Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

### EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

All experiments were conducted in accordance with the National Institutes of Health Guide for the Care and Use of Laboratory Animals (NIH Publication No. 80–23) and were approved by the Institutional Animal Care and Use Committee of the University of Iowa.

As described previously,[6] the data used in this study were collected from Sprague-Dawley Norway rats (*Rattus norvegicus*) at P8–9 (hereafter P8; n = 8) and P12–13 (hereafter P12; n = 8). Equal numbers of males and females were used, and all subjects were selected from different litters. See our previous report for additional information.

### METHOD DETAILS

As described previously,[6] high-speed (100 fps) video and M1 extracellular unit activity were collected from unanesthetized pups as they cycled through sleep and wake. Forelimb displacement was tracked using DeepLabCut;[66,67] discrete movements were identified with custom software and confirmed visually. Note that both twitches and wake movements exhibited similar distributions of limb displacement at P8 and P12 (see Figure 3A in Glanz et al., 2021).[6] See our previous report for additional information related to surgeries, electrophysiological recordings, video data collection, histology, spike sorting, and determination of behavioral state and forelimb movements.

## QUANTIFICATION AND STATISTICAL ANALYSIS

### Data preparation

For all decoding procedures, M1 unit activity and forelimb displacement served as the predictor and target variables, respectively. Both variables were binned in 20-ms increments. Forelimb displacement was calculated as the absolute value of zero-centered limb position along the x-, y-, and z-dimensions and was smoothed using a 20-ms half-width Gaussian kernel.

To prevent model overfitting, the 1-h recordings were split into training (36 min; 60%) and testing (9 min; 15%) datasets that were used, respectively, to train and test model parameters for best performance. To ensure an unbiased final assessment of decoder performance,[25] a validation dataset (15 min; 25%) was held back until the final model parameters were established; the validation dataset was used for the results reported here. For a tutorial on similar neural decoding procedures, see Glaser et al.[25]

### Data scaling

M1 activity was z-scored prior to decoding. The scaling factors (mean and standard deviation) were calculated using only the training dataset to avoid data leakage (i.e., positive bias in performance estimates due to information from the testing dataset leaking into the training dataset).[68] Forelimb position was normalized before decoding. The scaling factors (minimum and maximum) were again calculated using only the training dataset to avoid data leakage.

### Linear model

The linear model was composed of a single-layer Tensorflow model[69] with no hidden layer or activation function. This design forms a linear perceptron[70] and is equivalent to an ordinary least-squares (i.e., linear) regression. All models were assembled in Keras, a software package that assists in building Tensorflow models.[71] A neural-network approach was selected instead of a standard linear regression (which guarantees an optimal result) to match our nonlinear analyses. Using a standard linear regression did not meaningfully improve decoder performance (average $r^2$ improvement of 0.03). Finally, a 240-ms time window surrounding each timepoint was flattened (i.e., crossed with the feature of individual units) to provide additional temporal context to the model.

The decoder's learning parameters were as follows: Mean-square error was selected as the loss function (i.e., a function used to determine the error between the real limb displacement and the predicted limb displacement). Adam[72] was selected as an optimizer (i.e., an algorithm that updates the internal parameters of the model across training iterations). The learning rate (i.e., the degree to which the internal parameters are updated) was set to 0.001 arbitrary units, which is the default learning rate for Adam in the Keras package. We tested a variety of different combinations of loss functions, optimizers, and learning rates without observing a meaningful impact on decoder performance (average $r^2$ improvement of 0.02).

The model was allowed to train until the test loss function stopped decreasing, which typically occurred after 3–5 iterations. For each time bin, the decoder model predicted three target variables corresponding to x-, y-, and z-displacement of the forelimb. Before comparing predicted with actual limb displacements, the three predicted displacement variables were combined into a single variable using the Pythagorean theorem:

$$\sqrt{x^2+y^2+z^2} \quad \text{(Equation 1)}$$

Finally, we note that decoder performance was similar when each of the three dimensions was analyzed independently (average $r^2$ improvement of 0.02).

### Nonlinear model

Several nonlinear models were compared before choosing a long-short-term memory (LSTM) neural network.[26] The LSTM model was chosen for its superior performance on the current dataset and its successful application in similar neural decoding applications.[25,73,74] Like all models with many parameters, LSTMs are sensitive to initialization parameters and are prone to overfitting. We addressed these limitations by repeatedly randomizing the initialization parameters and by performing all analyses on a held-out validation dataset.

The number of nodes in the input layer was identical to the number of units in each specific M1 recording (i.e., 14–38 nodes). Models with a larger numbers of nodes were tested (up to 500) but resulted in only minor improvements in decoder performance (average $r^2$ improvement of 0.02). A 240-ms time window, similar to that used for the linear model, was included as the model's recurrent (i.e., time-based) feature.

To further prevent overfitting, a penalty was added to the recurrent dimension of the model to prevent rapid changes to the model's internal parameters across training iterations. The penalty selected was L2 (ridge regression[75]):

$$\lambda \cdot \|\beta\| \quad \text{(Equation 2)}$$

where $\lambda$, which represents the magnitude of the penalty, was set to 0.001 arbitrary units.

The optimizer, learning rate, and loss function were identical to the linear model. The nonlinear model was also allowed to run until the test loss stopped decreasing (typically within 3–5 iterations). Similar to the linear model, three independent predictions (x-, y-. and

z-displacement) were made for each time bin and reduced to a single variable using the Pythagorean theorem (Equation 1), after which predicted and actual limb displacements were compared.

### Decoder performance

Decoder performance was evaluated using the square of the Pearson correlation coefficient ($r^2$) with actual (x) and predicted (y) limb displacement as the two variables:

$$r^2 = \left( \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum (x_i - \bar{x})^2 \sum (y_i - \bar{y})^2}} \right)^2 \quad \text{(Equation 3)}$$

In typical neural decoding applications, regression performance is measured using the Coefficient of Determination ($R^2$)

$$R^2 = 1 - \frac{\sum (y_i - x_i)^2}{\sum (x_i - \bar{x})^2} \quad \text{(Equation 4)}$$

which is sensitive to the scale of the errors. The scale-invariant definition of model performance used here (Equation 3) solves the "intermittent demand" problem.[76] Briefly, because the limb's displacement is near zero (i.e., at rest) across the vast majority of sampled time points, $R^2$ disproportionately rewards predictions equal to the mean limb displacement and punishes predictions that vary from baseline, which is not the case with $r^2$.

### Model lesion experiments

Two "lesion" experiments were performed in which individual M1 units or action potentials (spikes) were randomly selected to be removed from the decoding process. A given percentage of units or spikes, over a range from 0% to 90% (in increments of 10%), was randomly selected and removed without replacement using the NumPy software package in Python.[77] The resulting units or spike trains were then used as the predictor variable and entered into the decoder model that was trained on the "intact" dataset. The performance of the decoding model after a lesion was calculated as the ratio of the lesion $r^2$ to the intact $r^2$, converted to a percentage. The randomization procedure was repeated ten times and the ten $r^2$ values were averaged for each pup before statistical comparison.

### Mutual information

Mutual information between a target unit ($X_0$) and a subset of predictor units ($\{X_k\}$) of size $k$ was defined as:

$$I(X_0; \{X_k\}) = S(X_0) - S(X_0 \mid \{X_k\}) \quad \text{(Equation 5)}$$

where $I(X_0; \{X_k\})$ represents the linear mutual information regarding the response of a target unit and of a subset of units[78–80] to a particular movement. This is equivalent to the entropy of the target unit ($S(X_0)$) minus the entropy of the target unit explained by the subset ($S(X_0 \mid \{X_k\})$). Mutual information was then calculated.[81] Random subsets of units were randomly selected from all possible combinations (10 selections per subset of size $k$).

### Model transplant experiment

To test the interchangeability of decoder models between pups of a particular age, a "model transplant" experiment was performed. The M1 activity (predictor variable) of an individual rat (the "recipient") was used to predict forelimb displacement using the model weights from a different but same-aged pup (the "donor"). Two pups of the same age were paired when the recipient had at least the same number of M1 units as the donor. The two animals' units were then matched by descending firing rate. Any excess units from the recipient pup were discarded from the analysis. In total, 29 such donor-recipient pairs were selected for both P8 and P12 animals. To test whether firing-rate sorting introduced bias, the experiment was repeated 30 times with random unit-unit pairings. The performance of the "model transplant" was measured as the ratio of the transplanted $r^2$ (recipient × donor) to the original $r^2$ (recipient × recipient).

### Statistical analyses

Before statistical analysis, all data were tested for normality using the Shapiro-Wilk test, for equal variance using Levene's test (for between-subjects variables), and for sphericity using Mauchly's test (for within-subjects variables with >2 groups). For analyses in which the variance between groups was not equal, a pooled error term was not used when generating simple main effects and post-hoc tests. For analyses in which sphericity was violated, a Huynh-Feldt correction was applied to the degrees of freedom. All $r^2$ values were arc-sin transformed before analysis. The mean and standard error of the mean (SEM) were used throughout as measures of central tendency and dispersion, respectively.

All analyses were performed as independent t tests or two-way mixed-design ANOVAs. Simple main effects were only tested if the interaction term was significant. In all two-way ANOVAs, an adjusted partial eta-squared was used as an estimate of effect size that corrects for positive bias due to sampling variability.[82] For t tests, an adjusted eta-squared estimate of effect size was reported.

**Cell Reports, Volume *42***

## Supplemental information

# Neural decoding reveals specialized kinematic tuning after an abrupt cortical transition

**Ryan M. Glanz, Greta Sokoloff, and Mark S. Blumberg**

# Supplemental Information

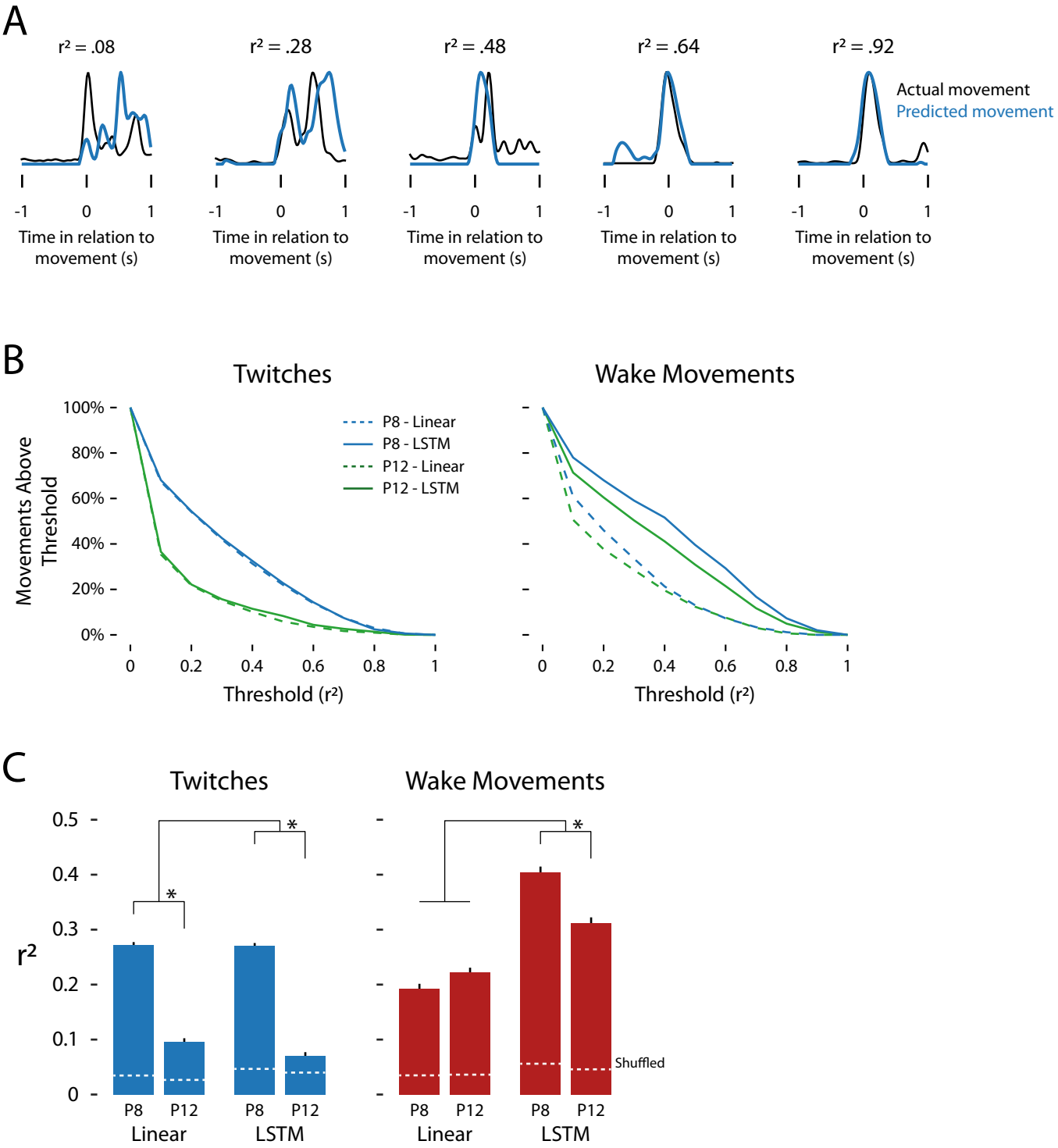



**Figure S1.** Decoder performance for twitches and wake movements.

(**A**) From left to right, actual forelimb displacement (black lines) and predicted forelimb displacement (blue lines) for a single representative twitch at five levels of decoder performance (from weak to strong). (**B**) Left: Plot depicting the percentage of twitches above a particular threshold of decoder performance (x-axis) for P8 (blue lines) and P12 (green lines) rats. Dashed and solid lines represent linear and nonlinear decoder performance, respectively. Right: Same as left, but for wake movements. (**C**) Same as Figure 2B, but after decoding was performed separately for sleep and wake periods. Note that the linear decoding of wake movements at P12 was more accurate than in Figure 2B, likely due to the discarding of sleep-related variance from this decoding process. Brackets denote a significant interaction between age and decoder ($p < .05$). Asterisks denote significantly better performance of the linear decoder at P8 compared to P12 ($p < .05$).