Article

Current Biology

Implicit mechanisms of intention

Highlights

- PPC encodes movement intent before the urge to move
- Subconscious neural dynamics emerge when participants choose to perform a task
- Our results suggest subconcious processing connects goals with goal-achieving actions
- Early coding drives untimely BMI actions that can be corrected with a decoder design

Authors

Tyson Aflalo, Carey Zhang, Boris Revechkis, Emily Rosario, Nader Pouratian, Richard A. Andersen

Correspondence

tyson.aflalo@gmail.com

In brief

High-level cortex encodes motor decisions before awareness, suggesting that choice is preconsciously determined. Aflalo et al. show that preconscious activity is triggered by the choice to participate in the experiment, thus fulfilling, not predetermining, choice. Associated neural dynamics require a decoder design to align the BMI output with conscious choice.

Aflalo et al., 2022, Current Biology *32*, 1–10
May 9, 2022
https://doi.org/10.1016/j.cub.2022.03.047

CellPress

Current Biology

CellPress
OPEN ACCESS

Article

Implicit mechanisms of intention

Tyson Aflalo,[1,2,5,*] Carey Zhang,[1] Boris Revechkis,[1] Emily Rosario,[4] Nader Pouratian,[3] and Richard A. Andersen[1,2]
[1]California Institute of Technology, Division of Biology and Biological Engineering, 1200 E California Blvd., Pasadena, CA 91125, USA
[2]California Institute of Technology, Tianqiao and Chrissy Chen Brain-Machine Interface Center, 1200 E California Blvd., Pasadena, CA 91125, USA
[3]University of California, Los Angeles, Geffen School of Medicine, 10833 Le Conte Ave, Los Angeles, CA 90095, USA
[4]Casa Colina Hospital and Centers for Rehabilitation, 255 E Bonita Ave, Pomona, CA 91767, USA
[5]Lead contact
*Correspondence: tyson.aflalo@gmail.com
https://doi.org/10.1016/j.cub.2022.03.047

SUMMARY

High-level cortical regions encode motor decisions before or even absent awareness, suggesting that neural processes predetermine behavior before conscious choice. Such early neural encoding challenges popular conceptions of human agency. It also raises fundamental questions for brain-machine interfaces (BMIs) that traditionally assume that neural activity reflects the user's conscious intentions. Here, we study the timing of human posterior parietal cortex single-neuron activity recorded from implanted microelectrode arrays relative to the explicit urge to initiate movement. Participants were free to choose when to move, whether to move, and what to move, and they retrospectively reported the time they felt the urge to move. We replicate prior studies by showing that posterior parietal cortex (PPC) neural activity sharply rises hundreds of milliseconds before the reported urge. However, we find that this "preconscious" activity is part of a dynamic neural population response that initiates much earlier, when the participant first chooses to perform the task. Together with details of neural timing, our results suggest that PPC encodes an internal model of the motor planning network that transforms high-level task objectives into appropriate motor behavior. These new data challenge traditional interpretations of early neural activity and offer a more holistic perspective on the interplay between choice, behavior, and their neural underpinnings. Our results have important implications for translating BMIs into more complex real-world environments. We find that early neural dynamics are sufficient to drive BMI movements before the participant intends to initiate movement. Appropriate algorithms ensure that BMI movements align with the subject's awareness of choice.

INTRODUCTION

A neural prosthesis translates the motor intentions of paralyzed individuals into control signals for assistive devices.[1–4] These systems have largely assumed that decoded neural signals reflect the conscious intentions of the user. However, high-level cortical regions can encode intended actions before conscious awareness or absent awareness.[5–9] The prospect of a neural prosthetic interface executing movements before or without the explicit awareness of the individual raises legal, ethical, and practical challenges for brain-machine interface (BMI) adoption. Despite this, the origins of early "preconscious" activity and how this activity might interact with a neural prosthesis have not been explored.

The posterior parietal cortex (PPC) is an important region at the intersection of action awareness and production as intentions are encoded by early planning activity,[3,10–14] stimulation can induce an urge to initiate movement,[15] and damage can disrupt awareness of movement and body state.[16–19] The significance of this body of literature for neural prosthetics was highlighted when, as part of an ongoing clinical study, we developed a neural prosthetic piano interface using signals from PPC (Video S1). We trained an algorithm to decode individuated finger movements in response to simple cues and mapped decoded movements to the keys of a virtual piano. However, while playing the piano, the participant sometimes reported that "the keys just automatically hit themselves without me thinking about it. It just seemed like it (the decoding algorithm) knew the tune and did it on its own. I did not have to move my fingers to make it happen" (Video S2). This example highlights how neural prosthetic systems decoding neural signals from high-level regions of cortex can execute actions before the user explicitly intends movement.

Here, we study the timing and nature of single neuron and population activity in the PPC of two tetraplegic individuals participating in a human neural prosthetic clinical study. Participants chose when to move and retrospectively reported the time of the choice. In two follow-up experiments, a participant chose whether to participate in the task on a trial-by-trial basis or chose both when to move and what movement to perform. Results from these experiments suggest that within our recording regions of PPC (1) the rapid rise in neural activity before movement awareness and initiation are part of network dynamics associated with movement planning, (2) cortical activity is better correlated with motor production than awareness, (3) cortical responses are contingent on high-level (nonmotor) choices, (4) cortical activity

CellPress
OPEN ACCESS

Current Biology
Article

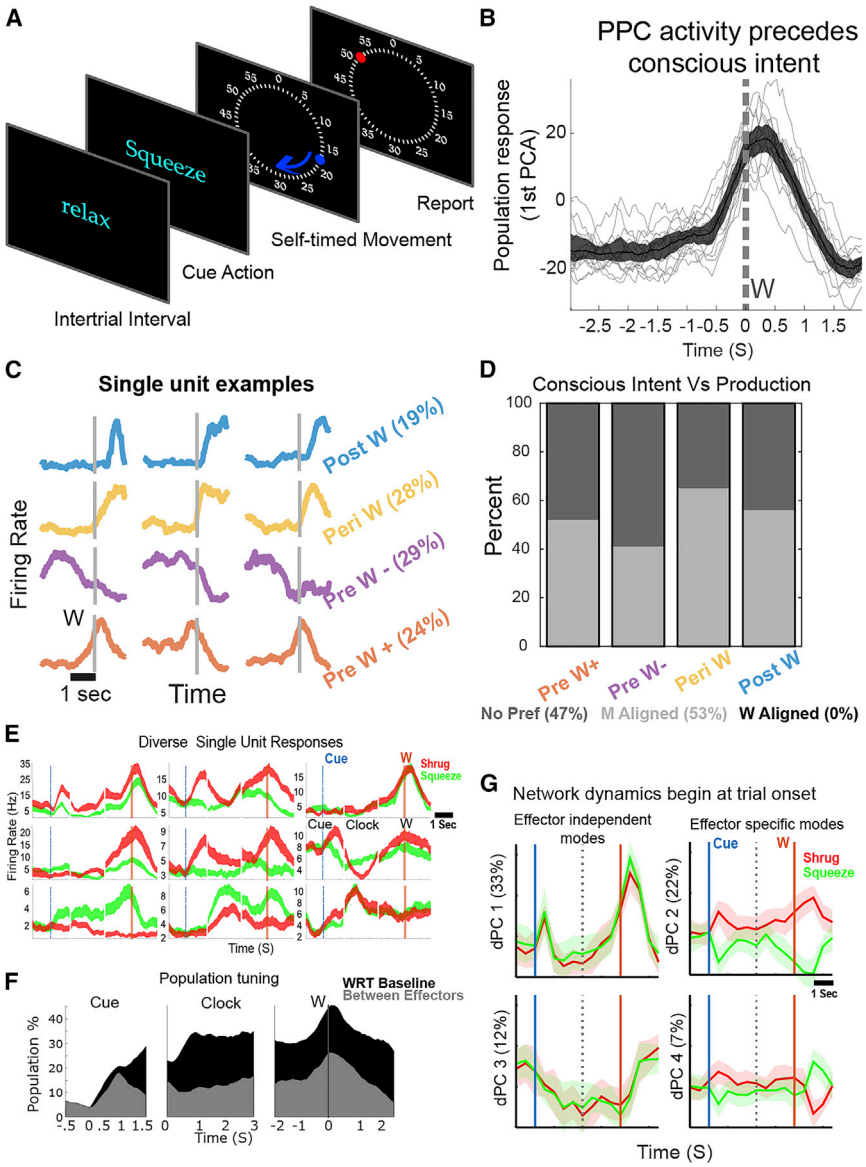

Figure 1. Neural dynamics demonstrate early encoding of movement intent

(A) Task paradigm.

(B) Population activity from a representative session (participant NS) summarized as 1st principal component of population response (42% variance explained, mean ± SEM with single-trial examples of the population response in gray; 200 ms boxcar smoothing).

(C) Single unit examples illustrating diverse temporal responses (500 ms boxcar smoothing). Colors identify four basic temporal profiles found within the population (cluster analysis, Bayesian information criteria to determine the number of clusters). Percent of total population falling into each cluster shown in parenthesis.

(D) Proportion of neurons whose temporal response is best explained relative to reported urge to move (W aligned), EMG onset (M aligned), or neither (no preference), broken up by cluster identity. Percentages in legend (bottom) refer to the total percentage of population collapsing across neural classes identified in (C).

(E) Sample neural responses illustrating effector specific and effector general dynamics beginning with trial onset (mean ± SEM). Each panel illustrates a separate unit. Panels from left to right are aligned to cue onset, clock onset, and time of reported urge (W).

(F) Percent of the population demonstrating significant modulation ($p < 0.05$ uncorrected, linear regression) from baseline (black) and significant differences between effectors (gray) through trial progression.

(G) Population-level latent dimensions demonstrating effector independent and specific network dynamics (cross-validated mean ± 95% CI). The dashed line represents temporal discontinuity from concatenating cue-aligned and movement-aligned signals. We adopt the concatenated visualization for supervised learning techniques to emphasize dependencies between time points.

See also Figures S1–S3.

can occur without awareness of motor intent when in line with high-level choices, and (5) algorithm selection and design are essential to ensure that high-level motor intention signals are appropriately translated into prosthetic control signals. One interpretation is that PPC contributes to an internal model of the motor planning network that transforms high-level task goals into the motor commands necessary to achieve those goals. Further, our results challenge traditional interpretations of ''preconscious'' activity by demonstrating that neural population responses that rise before the participant's choice to initiate movement first arise due to the participant's choice to perform the task.

RESULTS

We recorded single-neuron activity in two tetraplegic individuals (N.S. and E.S.) implanted with Neuroport arrays near the junction of the postcentral and intraparietal sulci in the PPC (Figure S1). We used three complementary paradigms in which participants performed self-initiated movements to understand the relationship between conscious intent and neural activation. N.S. and E.S. participated in experiment 1. N.S. participated in experiments 2 and 3.

The timing of single-neuron activity during motor production and awareness

In the first experiment (Figure 1A), each trial began with a cue instructing one of two movements of the contralateral upper limb: brief shrugs of the shoulder (a movement above the spinal cord injury) and attempted squeezing of the hand (a movement below the spinal cord injury). The participant was then free to choose when they initiated the instructed movement (Figure S2). We used the method of Libet[6] to allow participants to retrospectively report when they first experienced the urge to initiate movement.

Please cite this article in press as: Aflalo et al., Implicit mechanisms of intention, Current Biology (2022), https://doi.org/10.1016/j.cub.2022.03.047

Following Libet,[6] we refer to this reported time as "W" and the time of the motor response as "M."

PPC population activity increases before W, supporting previous findings that changes in neural activity precede awareness of intent[5–9] and extending these results to human PPC (Figure 1B). Further, the sensitivity to brain state provided by simultaneously recorded neural populations reveals that the early ramping activity seen in trial averages provides an accurate template for the population response measured during single trials (compare black and light gray traces of Figure 1B). However, individual neuron responses were diverse. Cluster analysis identified subpopulations becoming engaged before W (Pre W±) but also around the time of W (Peri W) or after (Post W) (Figure 1C). This cascade of neural activations in time points to complicated neural mechanisms underlying the process of voluntary movement initiation and is consistent with delayed-movement experiments in nonhuman primates (NHPs)[20–22] but previously unreported for voluntary movement initiation in humans.

Conscious intent and movement initiation are two events confounded in time. Early rising activity (Figure 1C [bottom row]) has been hypothesized to trigger the urge to initiate movement as activity passes a threshold.[7,23] However, this activity may reflect a parallel process, whereby movement is prepared for execution without a direct relation to awareness (e.g., movements can be performed without a concomitant awareness).[24,25] To address whether neural signals are more closely related to motor production or subjective experience, we looked at the relative timing of these signals. We can address this question because the time intervals between M and W varied (Figure S2B), with a range comparable to simple reaction time tasks.[26] We used the following logic: if a neural signal is time locked to motor production (M), then aligning neural signals to the time of conscious intent (W) will misalign the temporal profiles leading to greater trial-to-trial variability and vice versa (Figure S3). All neuron types were better explained in relation to the time of motor initiation or demonstrated no preference (Figure 1D). Neurons showing no preference had significantly lower signal to noise (depth of modulation/variance) and thus likely reflect detection difficulty instead of a mixture of W and M alignment. These results cannot be easily explained by increased measurement error when using the clock to report onset times (Figure S3). Together, our results suggest that early rising activity better reflects mechanisms related to motor production and not awareness as such.

What are the origins of neural activity changes that precede awareness?

In prior studies, activity preceding movement awareness is described as having "preconscious" origins; however, this label does not help us understand how systems of neurons give rise to behavior and experience. From a systems neuroscience perspective, the increase in neural activity must be the consequence of preceding neural computations (e.g., related to cue processing, context, high-level intent). Our motivation for including the effector cue was to introduce a known moment when movement-related information could first be encoded in the cortex. We hypothesized that analysis of neural activity at this moment would help contextualize activity arising around W.

Figure 1E shows neural responses from the effector cue through movement production for both shrug and squeeze trials. As above, many neurons exhibit rapid changes in activity hundreds of milliseconds before W (right portion of each subpanel). However, tracing the neural responses backward in time shows that rapid changes at the end of the trial are connected to temporal or effector-specific modulation beginning at cue presentation (Figures 1E and 1F). We used a cross-validated variant of demixed principal component analysis (dPCA),[27] a supervised dimensionality reduction technique, to analyze the temporal evolution of PPC activity at the population level (Figure 1G). These population responses are presumed to capture the underlying latent variables that define a cortical region.[28–30] These principal components demonstrate that the neural dynamics that precede the urge to initiate movement occurs within a latent neural subspace that is first engaged immediately after the movement is encoded (e.g., in each panel, latent population activity first modulates directly following the cue and then modulates further, often in anticipation of W). Early task-relevant modulation suggests that neural modulation around W is a consequence of neural processes initiated at (or before) trial onset and thus are the consequence of the participants' decision to perform the task as instructed.

The latent dimensions (Figure 1G) contain an effector-specific component (demixed components 2&4; consistent with an evolving motor plan) and an effector-general component (demixed compoents 1&3; potentially encoding global urgency, movement likelihood, timing intervals, or other effector-independent quantities), again drawing parallels with work in NHPs.[31–33] Both effector-specific and general components emerged in anticipation of W.

Is early activity the consequence of high-level intent?

To test the hypothesis that neural activity before W is a consequence of the participant's decision to perform the task, we had participant N.S. choose, on a trial-by-trial basis, whether to perform a trial ("opt out task"; Figure 2A). We wanted the opt out decision to occur before the effector cue (e.g., before the subject knew which effector to move) to dissociate the high-level intent to perform the trial from motor planning signals. Therefore, the participant was instructed to abort the trial if the opt out choice was not made before a beep played 1.5 s before the effector cue. We reasoned that if neural dynamics are the consequence of NS's high-level choice, then neural dynamics would only occur during trials N.S. chose to perform.

N.S. aborted one trial (total trials = 286) because she had not chosen to participate by the beep. When N.S. chose to perform a trial (65.2% of trials), we measured selective neural responses consistent with task objectives at the time of the effector cue; conversely, this activity was not encoded or quickly suppressed when N.S. chose to forgo a trial (Figures 2B and 2C). Thus, the observed neural dynamics are a consequence of the participant's choice to perform the task. We did not find neural encoding of the choice to perform a given trial before the effector cue (Figure 2D).

Are early planning dynamics a byproduct of the task paradigm?

What is the timing of motor planning signals when subjects are free to choose what to move in the absence of task structure? In the above paradigms, the effector cue provides a known

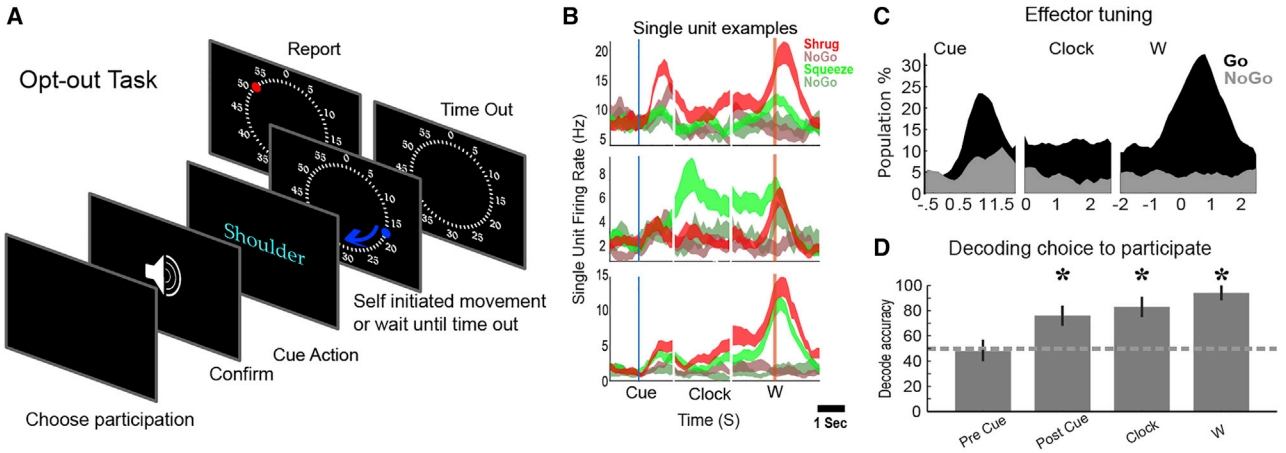

Figure 2. Volitional neural dynamics connect trial-onset to movement production

(A) Task paradigm.

(B) Single unit examples illustrate how neural behavior depends on the participant's choice (N trials = 20 ± 4, mean ± SEM).

(C) Percent of the population ($p < 0.05$ uncorrected, linear regression) exhibiting differential modulation to effector cue contingent on high-level response.

(D) Population decoding of the participant's choice to participate on a trial-to-trial basis (cross-validated mean ± 95% CI). Asterix indicates significant decoding (shuffle test.).

moment when the motor plan is first encoded. However, this explicit external cue may create neural behavior that otherwise would not occur in more typical circumstances, such as if the subject was free to choose which effector to move. It also promotes awareness of the cued effector. To mitigate against these limitations, we asked participant N.S. to choose which of three movements to perform and the timing of the movement within a simple task environment (Figure 3A; see also Stetson and Andersen[13]). N.S. was given minimal instruction; however, she was asked to (1) initiate movement immediately upon deciding which movement to perform and (2) to "avoid repeated sequences of movements" such that the movements were "spontaneous and unpredictable." To allow for a more natural, self-paced flow of movements, we did not require N.S. to report the time of the experienced choice. The goal of this task variant was 2-fold: first, to test whether rising activity just before movement would also be observed in more natural testing conditions, and two, to measure when intention-related activity emerges when not explicitly cued.

Intervals between movements were broadly distributed (Figure 3B). Movement choices were well balanced across the available options, and the transitions between options were well sampled (Figure 3C). The behavior of individual neurons demonstrated diverse and complex temporal patterns (Figure 3D) consistent with the earlier structured tasks (e.g., Figures 1C and 1E). To understand timing, we aligned neural data to movement onset and used classification analyses to decode both the current executed movement and the next movement. Data were median split into long and short intermovement intervals (Figure 3B [median = 1.9 s]) and analyzed separately to understand the influence of movement intervals. In line with the results discussed above, information about the current movement begins to rise sharply ~1 s before initiation (Figure 3E [blue]). Information about the next movement began nearly coincident with the execution of the current movement (−130 ± 90 ms for short-interval trials and 56 ± 90 ms for long-interval trials, Figure 3E [red]). In the case of long interval trials, this early coding occurs a minimum of 1.9 s before movement initiation, nearly a second before the ramping responses in the instructed paradigms above (Figures 1 and 2). Such early encoding suggests that implicit neural computations support behavior when the subject is free to choose both the time and type of action.

In theory, statistical regularities in transitions between actions could drive significant decoding of the forthcoming movement based on neural coding of the current movement. To address this possibility, we repeated the classification analysis (Figure 3E) after regressing out activity attributable to the current action. The pattern of results remained the same (−114 ± 100 ms for short-interval trials and 32 ± 90 ms for long-interval trials).

Early activity and consequences for neural decoding

Can early engagement of the motor planning network explain decoder activation before the explicit intention to initiate movement? We looked at how population-level activity collected in experiment 1 (Figure 1) dynamically evolved in a space directly relevant to neural classification. We computed the normalized Mahalanobis distance between neural population activity recorded through the course of a trial in relation to neural activity recorded during the intertrial interval (ITI) and movement execution (Go) epochs (Figures 4A and 4B). Relative distances in Mahalanobis space underlies classification using linear discriminant analysis and can thus directly measure the factors that drive classification. We focused our analysis on shoulder trials to ensure the absence of electromyogram (EMG) activity before W (therefore, any activity before W must reflect internal processes and not overt movement). Directly after the cue, we measure a phasic response toward the population state associated with the Go epoch, presumably reflecting the encoding of the motor plan. The neural population then reaches a steady state, on average, holding a position approximately midway to the decision boundary that separates neural states associated with the ITI and Go epochs. Next, before movement, the population drifts

Current Biology
Article

CellPress
OPEN ACCESS

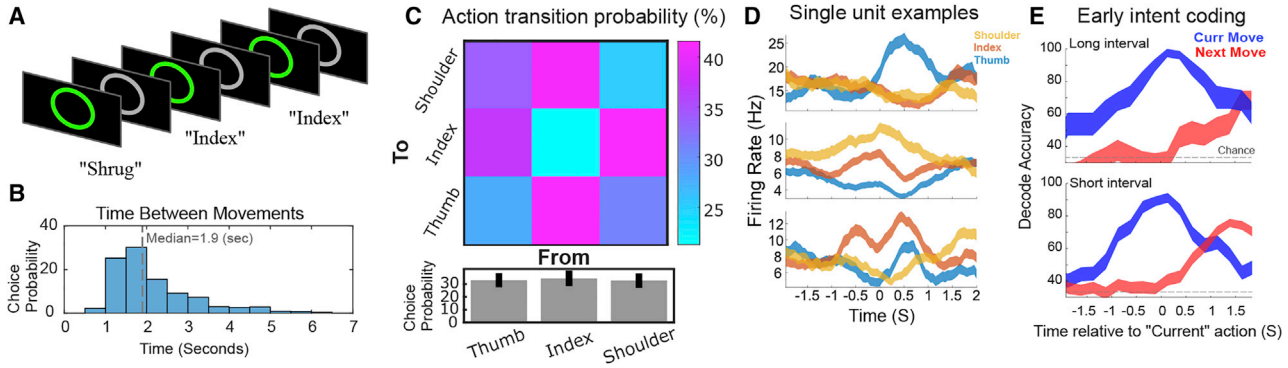

Figure 3. Early coding of motor intentions in a simplified choice task

(A) Task. The subject was free to move the shoulder or attempt movement of the thumb or index whenever she felt the urge to do so. There was no task structure outside the brief (250 ms) change of annulus color from gray to green immediately after the movement had been detected following movement onset.

(B) Histogram of intervals between voluntarily initiated movements.

(C) Behavior during the free-choice task. Top: transition probabilities between actions. Bottom: percent of trials each action was performed (mean ± SEM computed across sessions).

(D) Single unit examples aligned to the verbal report of current movement onset (mean ± SEM).

(E) Accuracy decoding current and next action in the voluntary movement sequence split by interval duration (mean ± SEM across six sessions).

toward the execution response until approximately 1 s before movement initiation, at which point the ramping described above occurs (Figure 4C). We show cross-validated classification analysis of the same data in Figure 4D. The described neural dynamics can drive early decoding before the explicit intention to initiate action. This early decoding does not reflect indiscriminate errors rooted in noisy data; instead, early decoding reflected the network dynamics associated with a specific motor plan. We find that more sophisticated decoding approaches, in our case carefully structured hidden Markov models informed by the neural dynamics of our task, align the output of the neural decoder with the explicit timing of the user's intent (STAR Methods section improving classification; Figures 4D and S4; Video S3).

DISCUSSION

Several studies in human subjects have established that changes in neural activity occur before awareness of the urge to initiate movement.[5–9,18] These results are compelling and controversial because they suggest that unconscious neural processes predetermine behavior before an individual consciously processes the decision.[5] However, our results suggest that many examples of "preconscious" activity may instead reflect the mechanistic process by which high-level goals are transformed into the motor commands that achieve those goals. In this view, the pivotal decision occurs when the brain self-configured to perform a movement in the future.[14,34–36] It is for this reason that we have adopted the word "implicit" as opposed to "preconscious" when referencing neural responses immediately preceding "W." In other words, using the term "preconscious" submits to the logic that the neural events of consequence immediately precede movement initiation. However, taking a more holistic view, the activity before "W" may be better described as post- or peri-conscious as it comes after and is the consequence of the decision to perform the task. This implicit activity is analogous to our motor system. The conscious intent to drink a cup of coffee is automatically transformed into the spatiotemporal muscle activation patterns necessary to reach and grasp. Just as we are unaware of the neural processes that produce appropriate muscle activations, these data indicate that we are unaware of the internal details of the neural computations that support the production of future behaviors.

We refer to activity preceding the urge to initiate a movement as implicit; however, the participants are likely aware of which movements they will soon perform in experiments 1 and 2 (Figures 1 and 2). What is presumed implicit is the mechanistic process by which high-level task rules combine with environmental information (e.g., task cues) to produce movements at variable times in the future. The complex and diverse dynamics we see at the single neuron level (e.g., Figure 1E) likely play a role in this transformation (see below), but how these dynamics generate behavior is beyond our awareness. Further, experiment 3 (Figure 3) suggests that neural responses can be fully implicit. In the simple task, long-interval trials (those occurring greater than 1.9 s) occur outside the 1-s window that characterizes ramping activity before awareness in the Libet paradigm. During the interval of time when the motor plan transforms into movement, a new motor plan for the forthcoming movement is encoded into the planning network, and this activity occurs in the temporal window presumed to be outside awareness. However, this activity is still an expression of the participant's will to perform the task. Early encoding in the simple task is similar to NHP neural activity during well-rehearsed motor sequences.[37–40] Taken together, one speculation is that early implicit encoding of future motor acts may explain why N.S. felt the notes played themselves in the piano task.

PPC may encode the internal state of the motor production network

We analyzed trial-to-trial variability in the timing of neural signals relative to awareness ("W") and movement onset (as measured by EMG onset). We found that neural activity is better explained relative to movement onset, suggesting that our recorded neural

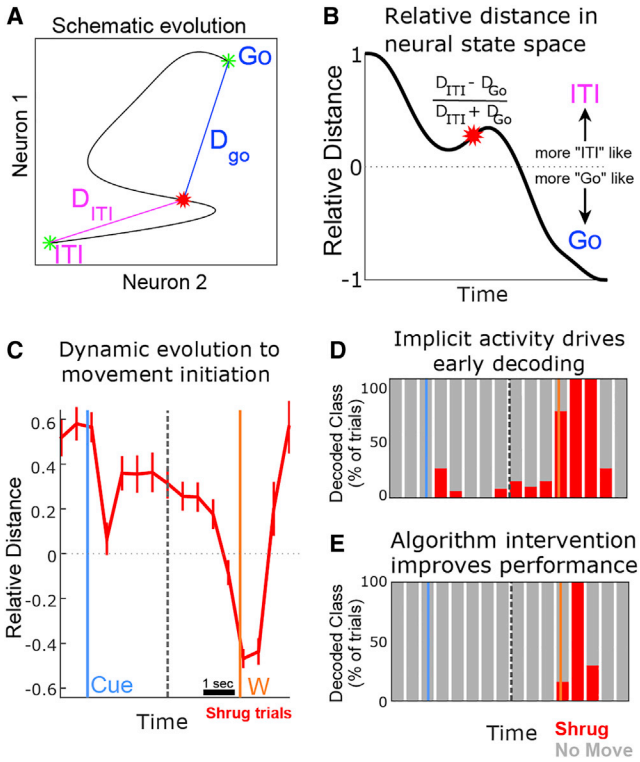

Figure 4. Early network dynamics can explain the "preconscious" triggering of the neural decoder

(A) Schematic. Through the course of a trial neural activity transitions from resting levels (ITI) to the response measured during execution (Go) (thin black line). This activity can be quantified by the geometric distance to the ITI and Go activity patterns (D_{GO}, D_{ITI}).

(B) Schematic. D_{GO} and D_{ITI} can be used to calculate a continuous measure of similarity to ITI and Go.

(C) Distance analysis applied to actual data for shrug (red) and squeeze (green) trials (mean ± 95% CI). The dashed line represents temporal discontinuity from concatenating cue-aligned and movement-aligned signals.

(D) Decode analysis using LDA applied to the same data as (C). Each column shows the percentage of trials each class was decoded for a single time bin (500 ms nonoverlapping bins). Populations dynamics observed in (C) are sufficient to generate early decodes of shrug actions (e.g., following cue and before reported urge).

(E) Decode analysis using a modified algorithm and training protocol applied to the same data as (D) can restrict decoding actions to the time of intended execution.

See also Figure S4 and Video S3.

population is involved in motor planning and execution but not directly tied to awareness. Although our control study (Figure S4) helps to strengthen this conclusion, the fact that temporal measurements were different in kind, e.g., clock monitoring versus EMG measurements, complicates a definitive conclusion on this point. However, an association with planning/execution but not awareness is consistent with and provides a deeper understanding of deficits following PPC damage.[19] In their paper, Sirigu et al. report that individuals with damage to PPC preserve an ability to initiate movement and report a conscious urge to initiate movement. However, patients retrospectively report that the urge to move is roughly coincident with movement onset, whereas control subjects report the urge occurring ~200 ms before movement onset, approximately when we see a steep ramping of PPC activity. The lesion study and our single unit results suggest that PPC encodes the internal state of the motor production network in a way that anticipates motor production, thus enabling early detection of a forthcoming movement. In other words, PPC participants cannot predict their own movements because they lack the anticipatory PPC neural activity reported in our study. This interpretation suggests that the previously described role of PPC in state estimation during movement[16,41–44] is broader than previously described and encompasses state estimation during planning formation through movement execution. Anticipatory signaling of the conversion of motor plans to motor initiation would have clear adaptive advantages. Environmental contexts are continually changing, and a movement programmed in the past may not be appropriate in an updated context. Anticipatory signals may ensure that a previously encoded motor plan remains relevant and allow an individual to cancel the movement otherwise. Anticipatory signals may also assist when assigning causal attribution to movement, e.g., to ensure that the brain knows the movement has an endogenous origin.[45]

We found that premovement PPC activity, which drives population neural activity closer to the classifier decision boundary (Figure 4; Video S3), can spuriously trigger a BMI classifier. This finding is seemingly at odds with studies in the monkey primary (M1) and premotor cortices. Kauffman et al.[22] report that planning activity is orthogonal to movement execution activity. This orthogonality is theorized to enable planning without generating overt behavior, the very problem we see in our classifier output. However, PPC is further removed from the motor output pathway and thus may employ different encoding strategies. For example, PPC may simplify telegraphing upcoming motor behaviors by providing similar output for planning and execution states. Other differences, such as species differences, the influence of long-standing injury, or methodological differences, may also contribute.

A node in the intention network

Although the Libet paradigm is presented as a single task, it includes two dissociable subtasks: a movement initiation task and a report task. The movement initiation task requires that the participant encode a motor program to initiate movement at variable and unpredictable times in the future. The reporting task requires that the participant create a "theory-of-mind" construct to report subjective experience. Our findings (that single-trial population activity arises 100 s of ms before W and neural timing best correlates to movement initiation rather than W) suggest that PPC neurons at the implanted location are related to the first task but not the second. Consistent with this view, neurons encode which movement the participant will perform before the participant is aware of the choice within a simple movement choice task (e.g., Figure 3).

The clear suggestion is that neurons responsible for constructing our subjective experience must rely on different populations of neurons located in other parts of the brain. This view is consistent with an emerging picture of the neural basis of consciousness and its neuroanatomical substrates. The PPC is a large and functionally heterogenous.[46,47] Our recording location within the anterior-superior aspect of PPC is primarily associated with

Please cite this article in press as: Aflalo et al., Implicit mechanisms of intention, Current Biology (2022), https://doi.org/10.1016/j.cub.2022.03.047

planning and monitoring body movements, as discussed above.[47] In contrast, posterior-inferior portions of PPC contain neural populations important for constructing our subjective experience of the world.[47–49] For example, electrical stimulation of the inferior parietal cortex induces the subjective experience of wanting to move or having moved, neuroimaging identifies neural correlates of awareness, and lesions disrupt awareness while otherwise preserving basic information processing. One possibility is that our recorded neural population provides inputs to inferior portions of PPC and thus contributes to, but is not directly responsible for, the participant's subjective experience of the task.

Our results are also consistent with a functional dissociation between high-level intent and motor planning. In recent work, we have found that PPC encodes a diverse array of variables related to movement intention and body state, including motor plans and trajectories, movements from diverse regions of the body, cognitive-motor strategy, observed actions, action semantics, and experienced and cognitive tactile sensations.[50–53] These results suggest that PPC may encode all task-related variables. However, somewhat to our surprise, data from the opt out task suggest that more anterior subregions of the PPC that we have implanted may have a limited role in representing more abstract, nonmotor forms of intent. In the opt out task, the subject was asked to choose, on a trial-by-trial basis, whether to perform the trial. This choice represents an example of a nonmotor decision because selecting a goal or desired outcome (e.g., I will or will not participate) occurs before knowledge of the specific motor act necessary to attain the goal (e.g., before the effector cue). We could not decode the decision before the effector cue, suggesting that our subregion of PPC does not encode nonmotor forms of intention. The separation of motor planning and nonmotor forms of intent is consistent with hierarchical intention models. In these models, the frontal cortex (e.g., prefrontal, orbitofrontal) or more posterior regions of PPC encode abstract decision variables, whereas areas closer to the central sulcus encode variables related to motor planning, body state, and movement execution.[54,55]

Relation to previous work on the early rising activity

There are two primary accounts for the early rise in neural activity that precedes movement onset and awareness. The original account of Libet assumes that a subconscious decision triggers movement preparation and initiation reflected in the early rising neural responses.[6] Critically, the decision is assumed to occur at the first detectable rise of the "readiness potential," a signature of spontaneously generated movements that appears 1 s before movement initiation. A second account, by Schurger et al.,[34] proposes a stochastic decision model by which neural activity undergoes a continuous random walk in the premovement period. Movement is quickly initiated, within ~150 ms, when the random walk crosses a decision threshold. The presumed timing of threshold crossing and reports of W is roughly coincident in time, indicating that the subjective experience of the decision occurs as the decision is made (not significantly earlier, as suggested by Libet). By this account, the early ramping activity is not the consequence of a decision but the final segment of the random walk that carries the neural state to the threshold.

Our results and interpretation are conceptually in agreement with the mechanistic formulation of Schurger et al. and subsequent studies by Murakami and colleagues.[34,56–58] In particular, these authors provide plausible accounts of how the high-level task goal of initiating movement at variable and unspecified times in the future can be implemented in neural populations. This mechanistic account is completely lacking from Libet's original formulation yet is necessary to fully appreciate how early rising activity relates to human agency. We add to this story by showing that these neural programs are instantiated based on the high-level goals of the participant. However, we find that the single-trial population-level activity of consistent temporal shape rises earlier than the 150 ms proposed by Schurger and colleagues.[34] Thus, our work is consistent with the idea that a distinct neural event leading to movement occurs hundreds of milliseconds before awareness. In our view, the mechanistic process leads to movement initiation, a process that precedes and is decoupled from awareness, as discussed above. Further, we find that the early neural activity is encoded as a motor plan, showing effector specific differences and not an abstract decision.[59,60]

Schurger et al. and Murakami et al. provide two separate mechanistic accounts for how movement timing is specified. In Schurger et al., single-trial timing results from different realizations of a stochastic leaky-accumulator model with identical initial conditions. In Murakami et al., the timing of future movements is specified at trial-onset by neurons that set the rate of the accumulator model. We did not find evidence that movement timing is determined at trial onset. However, our results may simply be a product of our recording region in the dorsal-anterior portion of PPC. For example, in a study in which monkeys made movements at different intervals in the future, dorsal-anterior parts of the monkey PPC looked similar to our recordings, although a more ventral-posterior portions of PPC showed varying rates of ramping beginning at trial onset.[35] Future studies recording from additional regions of the cortex are needed to address how humans program the timing of future actions.

Implications for neural prosthetics

When using a BMI to control computer devices, deviations of decoded control signals from the presumed intentions of the user are often attributed to the intrinsic noisiness of neural signals. However, recent work challenges this assumption by demonstrating that much of the neural fluctuations typically attributed to "noise" are correlated to often unmeasured subtle and high-dimensional features of the animal's behavior.[61,62] Especially in humans, such neural variability is likely also driven by internal processing related to conscious thoughts or subconscious processing. Understanding the sources and impact of these drivers of neural variability will be increasingly crucial as BMI systems are taken outside constrained laboratory experiments and tested in complex real-world environments. In our case, we find that early task-related neural dynamics are sufficient to drive BMI movements before the participant intends action.

We show that specific decoding approaches allow decoded output to reflect the participants' explicit motor intent accurately. Further, our results demonstrate that the relative timing of neural signals, explicit intent, and motor execution can become an algorithm design choice. Future work can explore how different

decoder settings within closed-loop experiments interact with the subjective experience of using a BMI, such as feelings of agency. Beyond questions of timing, behaviors that have been long rehearsed, such as typing, occur without awareness of the motor acts and require decoding implicit signals. In contrast, other unpracticed behaviors may require conscious control.[63] Both how neurons track our explicit intentions and the ways they do not are important to understanding and implementing a neural prosthetic system that enables effortless control across a range of environmental conditions.

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 SUBJECT DETAILS
- METHOD DETAILS
  - Data acquisition
  - Task descriptions
  - Analysis methods

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.cub.2022.03.047.

ACKNOWLEDGMENTS

We thank N.S. and E.S. for their bravery and hard work in making this work possible. This work was funded by NIH R01-EY015545, UG1-EY032039, the Della Martin Foundation, and the Boswell Foundation.

AUTHOR CONTRIBUTIONS

T.A. designed experiments with assistance from R.A.A., B.R., and C.Z. T.A., B.R., and C.Z. collected data. N.P. performed the surgery on N.S. T.A. analyzed the data. T.A. wrote the paper. T.A. and R.A.A. revised the paper.

DECLARATION OF INTERESTS

N.P. consults for Second Sight Medical Products and Abbott Laboratories. All other authors declare that they have no conflicts of interest.

Received: August 6, 2021
Revised: February 3, 2022
Accepted: March 15, 2022
Published: April 6, 2022

REFERENCES

1. Hochberg, L.R., Serruya, M.D., Friehs, G.M., Mukand, J.A., Saleh, M., Caplan, A.H., Branner, A., Chen, D., Penn, R.D., and Donoghue, J.P. (2006). Neuronal ensemble control of prosthetic devices by a human with tetraplegia. Nature *442*, 164–171.
2. Collinger, J.L., Wodlinger, B., Downey, J.E., Wang, W., Tyler-Kabara, E.C., Weber, D.J., McMorland, A.J., Velliste, M., Boninger, M.L., and Schwartz, A.B. (2013). High-performance neuroprosthetic control by an individual with tetraplegia. Lancet *381*, 557–564.
3. Aflalo, T., Kellis, S., Klaes, C., Lee, B., Shi, Y., Pejsa, K., Shanfield, K., Hayes-Jackson, S., Aisen, M., Heck, C., et al. (2015). Decoding motor imagery from the posterior parietal cortex of a tetraplegic human. Science *348*, 906–910.
4. Pandarinath, C., Gilja, V., Blabe, C.H., Nuyujukian, P., Sarma, A.A., Sorice, B.L., Eskandar, E.N., Hochberg, L.R., Henderson, J.M., and Shenoy, K.V. (2015). Neural population dynamics in human motor cortex during movements in people with ALS. Elife *4*, e07436.
5. Libet, B. (1985). Unconscious cerebral initiative and the role of conscious will in voluntary action. Behav. Brain Sci. *8*, 529–539.
6. Libet, B., Gleason, C.A., Wright, E.W., and Pearl, D.K. (1983). Time of conscious intention to act in relation to onset of cerebral-activity (readiness-potential)—the unconscious initiation of a freely voluntary act. Brain *106*, 623–642.
7. Fried, I., Mukamel, R., and Kreiman, G. (2011). Internally generated preactivation of single neurons in human medial frontal cortex predicts volition. Neuron *69*, 548–562.
8. Soon, C.S., Brass, M., Heinze, H.-J., and Haynes, J.-D. (2008). Unconscious determinants of free decisions in the human brain. Nat. Neurosci. *11*, 543–545.
9. Haggard, P., and Eimer, M. (1999). On the relation between brain potentials and the awareness of voluntary movements. Exp. Brain Res. *126*, 128–133.
10. Mountcastle, V.B., Lynch, J.C., Georgopoulos, A., Sakata, H., and Acuna, C. (1975). Posterior parietal association cortex of the monkey: command functions for operations within extrapersonal space. J. Neurophysiol. *38*, 871–908.
11. Andersen, R.A., and Cui, H. (2009). Intention, action planning, and decision making in parietal-frontal circuits. Neuron *63*, 568–583.
12. Cisek, P., and Kalaska, J.F. (2010). Neural mechanisms for interacting with a world full of action choices. Annu. Rev. Neurosci. *33*, 269–298.
13. Stetson, C., and Andersen, R.A. (2015). Early planning activity in frontal and parietal cortex in a simplified task. J. Neurophysiol. *113*, 3915–3922.
14. Maimon, G., and Assad, J.A. (2006). A cognitive signal for the proactive timing of action in macaque LIP. Nat. Neurosci. *9*, 948–955.
15. Desmurget, M., Reilly, K.T., Richard, N., Szathmari, A., Mottolese, C., and Sirigu, A. (2009). Movement intention after parietal cortex stimulation in humans. Science *324*, 811–813.
16. Wolpert, D.M., Goodbody, S.J., and Husain, M. (1998). Maintaining internal representations: the role of the human superior parietal lobe. Nat. Neurosci. *1*, 529–533.
17. Sirigu, A., Daprati, E., Pradat-Diehl, P., Franck, N., and Jeannerod, M. (1999). Perception of self-generated movement following left parietal lesion. Brain *122*, 1867–1874.
18. Desmurget, M., and Sirigu, A. (2009). A parietal-premotor network for movement intention and motor awareness. Trends Cogn. Sci. *13*, 411–419.
19. Sirigu, A., Daprati, E., Ciancia, S., Giraux, P., Nighoghossian, N., Posada, A., and Haggard, P. (2004). Altered awareness of voluntary action after damage to the parietal cortex. Nat. Neurosci. *7*, 80–84.
20. Churchland, M.M., and Shenoy, K.V. (2007). Temporal complexity and heterogeneity of single-neuron activity in premotor and motor cortex. J. Neurophysiol. *97*, 4235–4257.
21. Afshar, A., Santhanam, G., Yu, B.M., Ryu, S.I., Sahani, M., and Shenoy, K.V. (2011). Single-trial neural correlates of arm movement preparation. Neuron *71*, 555–564.
22. Kaufman, M.T., Churchland, M.M., Ryu, S.I., and Shenoy, K.V. (2014). Cortical activity in the null space: permitting preparation without movement. Nat. Neurosci. *17*, 440–448.
23. Schurger, A. (2018). Specific relationship between the shape of the readiness potential, subjective decision time, and waiting time predicted by an accumulator model with temporally autocorrelated input noise. eNeuro *5*, 0302–0317.2018.

24. Spering, M., and Carrasco, M. (2015). Acting without seeing: eye movements reveal visual processing without awareness. Trends Neurosci. *38*, 247–258.

25. Harms, I.M., Van Dijken, J.H., Brookhuis, K.A., and De Waard, D. (2019). Walking Without awareness. Front. Psychol. *10*, 1846.

26. Woods, D.L., Wyma, J.M., Yund, E.W., Herron, T.J., and Reed, B. (2015). Factors influencing the latency of simple reaction time. Front. Hum. Neurosci. *9*, 131.

27. Kobak, D., Brendel, W., Constantinidis, C., Feierstein, C.E., Kepecs, A., Mainen, Z.F., Romo, R., Qi, X.-L., Uchida, N., and Machens, C.K. (2014). Demixed principal component analysis of population activity in higher cortical areas reveals independent representation of task parameters. Preprint at arXiv, 1410.6031.

28. Saxena, S., and Cunningham, J.P. (2019). Towards the neural population doctrine. Curr. Opin. Neurobiol. *55*, 103–111.

29. Gallego, J.A., Perich, M.G., Miller, L.E., and Solla, S.A. (2017). Neural manifolds for the control of movement. Neuron *94*, 978–984.

30. Shenoy, K.V., Sahani, M., and Churchland, M.M. (2013). Cortical control of arm movements: a dynamical systems perspective. Annu. Rev. Neurosci. *36*, 337–359.

31. Kaufman, M.T., Seely, J.S., Sussillo, D., Ryu, S.I., Shenoy, K.V., and Churchland, M.M. (2016). The largest response component in the motor cortex reflects movement timing but not movement type. eNeuro *3*, ENEURO.0085-16.2016.

32. Churchland, A.K., Kiani, R., and Shadlen, M.N. (2008). Decision-making with multiple alternatives. Nat. Neurosci. *11*, 693–702.

33. Thura, D., and Cisek, P. (2014). Deliberation and commitment in the premotor and primary motor cortex during dynamic decision making. Neuron *81*, 1401–1416.

34. Schurger, A., Sitt, J.D., and Dehaene, S. (2012). An accumulator model for spontaneous neural activity prior to self-initiated movement. Proc. Natl. Acad. Sci. USA *109*, E2904–E2913.

35. Jazayeri, M., and Shadlen, M.N. (2015). A neural mechanism for sensing and reproducing a time interval. Curr. Biol. *25*, 2599–2609.

36. Romo, R., and Schultz, W. (1987). Neuronal activity preceding self-initiated or externally timed arm movements in area 6 of monkey cortex. Exp. Brain Res. *67*, 656–662.

37. Averbeck, B.B., Crowe, D.A., Chafee, M.V., and Georgopoulos, A.P. (2003). Neural activity in prefrontal cortex during copying geometrical shapes - II. Decoding shape segments from neural ensembles. Exp. Brain Res. *150*, 142–153.

38. Averbeck, B.B., Chafee, M.V., Crowe, D.A., and Georgopoulos, A.P. (2003). Neural activity in prefrontal cortex during copying geometrical shapes—I. single cells encode shape, sequence, and metric parameters. Exp. Brain Res. *150*, 127–141.

39. Genovesio, A., Brasted, P.J., and Wise, S.P. (2006). Representation of future and previous spatial goals by separate neural populations in prefrontal cortex. J. Neurosci. *26*, 7305–7316.

40. Averbeck, B.B., Chafee, M.V., Crowe, D.A., and Georgopoulos, A.P. (2002). Parallel processing of serial movements in prefrontal cortex. Proc. Natl. Acad. Sci. USA *99*, 13172–13177.

41. McNamee, D., and Wolpert, D.M. (2019). Internal models in biological control. Annu. Rev. Control Robot. Auton. Syst. *2*, 339–364.

42. Mulliken, G.H., Musallam, S., and Andersen, R.A. (2008). Forward estimation of movement state in posterior parietal cortex. Proc. Natl. Acad. Sci. USA *105*, 8170–8177.

43. Desmurget, M., and Grafton, S. (2000). Forward modeling allows feedback control for fast reaching movements. Trends Cogn. Sci. *4*, 423–431.

44. Graziano, M.S., Cooke, D.F., and Taylor, C.S. (2000). Coding the location of the arm by sight. Science *290*, 1782–1786.

45. Blakemore, S.J., Wolpert, D.M., and Frith, C.D. (1998). Central cancellation of self-produced tickle sensation. Nat. Neurosci. *1*, 635–640.

46. Andersen, R.A., Aflalo, T., Bashford, L., Bjånes, D., and Kellis, S. (2022). Exploring cognition with brain-machine interfaces. Annu. Rev. Psychol. *73*, 131–158.

47. Sirigu, A., and Desmurget, M. (2020). The sensorimotor posterior parietal cortex: From intention to action. In The Senses: a Comprehensive Reference (Elsevier), pp. 349–358.

48. Igelström, K.M., and Graziano, M.S.A. (2017). The inferior parietal lobule and temporoparietal junction: a network perspective. Neuropsychologia *105*, 70–83.

49. Webb, T.W., Igelström, K.M., Schurger, A., and Graziano, M.S.A. (2016). Cortical networks involved in visual awareness independent of visual attention. Proc. Natl. Acad. Sci. USA *113*, 13923–13928.

50. Chivukula, S., Zhang, C., Aflalo, T., Jafari, M., Pejsa, K., Pouratian, N., and Andersen, R.A. (2021). Neural encoding of felt and imagined touch within human posterior parietal cortex. Elife *10*, e61646. https://doi.org/10.7554/eLife.61646.

51. Aflalo, T., Zhang, C.Y., Rosario, E.R., Pouratian, N., Orban, G.A., and Andersen, R.A. (2020). A shared neural substrate for action verbs and observed actions in human posterior parietal cortex. Sci. Adv. *6*, eabb3984.

52. Zhang, C.Y., Aflalo, T., Revechkis, B., Rosario, E.R., Ouellette, D., Pouratian, N., and Andersen, R.A. (2017). Partially mixed selectivity in human posterior parietal association cortex. Neuron *95*, 697–708.e4.

53. Rutishauser, U., Aflalo, T., Rosario, E.R., Pouratian, N., and Andersen, R.A. (2018). Single-neuron representation of memory strength and recognition confidence in left human posterior parietal cortex. Neuron *97*, 209–220.e3.

54. Pacherie, E. (2008). The phenomenology of action: a conceptual framework. Cognition *107*, 179–217.

55. Haggard, P. (2008). Human volition: towards a neuroscience of will. Nat. Rev. Neurosci. *9*, 934–946.

56. Murakami, M., Shteingart, H., Loewenstein, Y., and Mainen, Z.F. (2017). Distinct sources of deterministic and stochastic components of action timing decisions in rodent frontal cortex. Neuron *94*, 908–919.e7.

57. Murakami, M., Vicente, M.I., Costa, G.M., and Mainen, Z.F. (2014). Neural antecedents of self-initiated actions in secondary motor cortex. Nat. Neurosci. *17*, 1574–1582.

58. Schurger, A., Hu, P.', Pak, J., and Roskies, A.L. (2021). What is the readiness potential? Trends Cogn. Sci. *25*, 558–570.

59. Travers, E., and Haggard, P. (2021). The Readiness Potential reflects the internal source of action, rather than decision uncertainty. Eur. J. Neurosci. *53*, 1533–1544.

60. Travers, E., Friedemann, M., and Haggard, P. (2021). The Readiness Potential reflects planning-based expectation, not uncertainty, in the timing of action. Cogn. Neurosci. *12*, 14–27.

61. 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.

62. 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.

63. Shadmehr, R., Smith, M.A., and Krakauer, J.W. (2010). Error correction, sensory prediction, and adaptation in motor control. Annu. Rev. Neurosci. *33*, 89–108.

64. Tibshirani, R., Walther, G., and Hastie, T. (2001). Estimating the number of clusters in a dataset via the gap statistic. J. R. Stat. Soc. *63*, 411–423.

65. Harris, K.D., Quiroga, R.Q., Freeman, J., and Smith, S.L. (2016). Improving data quality in neuronal population recordings. Nat. Neurosci. *19*, 1165–1174.

66. Cunningham, J.P., and Yu, B.M. (2014). Dimensionality reduction for large-scale neural recordings. Nat. Neurosci. *17*, 1500–1509.

67. Brendel, W., Romo, R., and Machens, C.K. (2011). Demixed Principal Component Analysis. Advances in Neural Information Processing Systems 24 (NIPS 2011).

68. Murphy, K. (2012). Machine Learning: A Probabilistic Perspective (The MIT Press).

69. Churchland, M.M., Cunningham, J.P., Kaufman, M.T., Foster, J.D., Nuyujukian, P., Ryu, S.I., and Shenoy, K.V. (2012). Neural population dynamics during reaching. Nature *487*, 51–56.

70. Michaels, J.A., Dann, B., Intveld, R.W., and Scherberger, H. (2015). Predicting reaction time from the neural state space of the premotor and parietal grasping network. J. Neurosci. *35*, 11415–11432.

71. Churchland, M.M., Cunningham, J.P., Kaufman, M.T., Ryu, S.I., and Shenoy, K.V. (2010). Cortical preparatory activity: representation of movement or first cog in a dynamical machine? Neuron *68*, 387–400.

72. Kemere, C., Santhanam, G., Yu, B.M., Afshar, A., Ryu, S.I., Meng, T.H., and Shenoy, K.V. (2008). Detecting neural-state transitions using hidden Markov models for motor cortical prostheses. J. Neurophysiol. *100*, 2441–2452.

Please cite this article in press as: Aflalo et al., Implicit mechanisms of intention, Current Biology (2022), https://doi.org/10.1016/j.cub.2022.03.047

CellPress
OPEN ACCESS

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Software and algorithms		
Matlab 2017a	Mathworks	https://www.mathworks.com
Other		
Neuroport system	Blackrock Microsystems	https://blackrockneurotech.com/

RESOURCE AVAILABILITY

Lead contact

Requests for further information and resources should be directed to and will be fulfilled by the lead contact, Tyson Aflalo (Tyson.aflalo@gmail.com).

Materials availability

This study did not generate new unique reagents.

Data and code availability

- The data that support the findings of this study are available from the lead contact upon reasonable request.
- This paper does not report original code, but our custom MATLAB code is available from the authors upon reasonable request.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Subject N.S. is a female human participant with a C3-C4 spinal lesion (motor complete). N.S. has no motor control or sensation below her upper trapezius. Subject E.G.S. is a male human participant with a C3-C4 spinal lesion (motor complete). E.G.S. has no motor control or sensation below his upper trapezius. All procedures were approved by the California Institute of Technology, University of California, Los Angeles, Rancho los Amigos National Rehabilitation Center, and Casa Colinas Centers for Rehabilitation Internal Review Boards. Informed consent was obtained from NS & EGS after the nature of the study and possible risks were explained. Consent to publish patient photos was obtained from NS and EGS. These studies were conducted after receiving permission from US Food and Drug Administration (Investigational Device Exemption), and data collected for this study is part of a registered clinical trial (additional information about the clinical trial is available at https://clinicaltrials.gov/ct2/show/NCT01958086). Study sessions occurred at Casa Colinas Centers for Rehabilitation and Rancho los Amigos National Rehabilitation Center.

METHOD DETAILS

Data acquisition

Behavioral setup

NS & EGS performed all tasks seated in their motorized wheelchair. Tasks were displayed on a 28 inch (NS) or 47 inch (EGS) LCD monitor. The monitors were positioned to occupy approximately 20 degrees of visual angle. Stimulus presentation was controlled using the Psychophysics Toolbox (23) for MATLAB.

Physiological recordings

Both NS and EGS were implanted with two 96-channel Neuroport arrays in putative homologues of area AIP and Brodmann's Area 5d.[3] Neural activity was amplified, digitized, and recorded at 30KHz with the Neuroport neural signal processor (NSP). The Neuroport System, comprising the arrays and NSP, has received FDA clearance for <30 days acute recordings. We received FDA IDE clearance (IDE #G120096, G120287) for extending the duration of the implant for purposes of a brain-machine interface clinical study using signals from posterior parietal cortex.

Unit activity was detected using thresholding at -3.5 times the root-mean-square after high-pass filtering (250Hz cut-off) the full-bandwidth signal. Single and multiunit activity was sorted using k-mediods clustering using the gap criteria[64] to determine the total number of neural clusters. Clustering was performed on the first n principal components where n was selected to account for 95% of

waveform variance. Results of offline sorting were reviewed and modified if the automated routine produced overly counterintuitive results following standard practice.[65] On average, 121 sorted units were recorded from NS per session and 29 from EGS per session.

Electromyogram (EMG) activity was recorded over the right trapezius muscle using Delsys EMG electrodes. Raw EMG activity was fed as an analog signal into the NSP and recorded time-locked to neural signals at a sampling rate of 2000 Hz. EMG signals were band-pass filtered (5th order Butterworth filter with cut-off frequencies of 5 and 250 Hz), full-wave rectified, and box-car smoothed using a 50 ms smoothing window.

Task descriptions

Modified Libet paradigm

We used the method introduced by Libet to enable the subject to retrospectively self-report the time they first experienced the urge to initiate a movement.[6] In a slight modification to the original paradigm, following an intertrial interval, the subject was cued to one of two possible actions (shrug of the contralateral shoulder or attempted squeezing of the contralateral hand) should be performed on a given trial (Figure 1A). Other aspects of the task were similar to that of Libet and subsequent studies,e.g.[5–7,9,19] An analog clock was presented with the initial position of the dial randomly (uniform distribution) selected. The dial rotated clockwise about the clock at 2.3 seconds per revolution. The subject was asked to withhold movement at least one complete clock cycle, at which point the participant was free to choose when to move. The instruction was given to "Keep your eyes on the clock but otherwise try to relax. When you first feel the urge to move, perform the movement, and note the position of the dial on the clock face when you first felt the urge to move. After you have completed the movement, let us know that you are done." At this point, the experimenter took control of the clock dial, positioned it according to the subject's reported time of awareness, and confirmed with the participant that this was the position of the dial when they first felt the urge to move. A minimum of 30 trials (15 trials per condition) were recorded per session. A total of 5 sessions in NS and two sessions in EGS were acquired. The distribution of time intervals the subjects withheld movement relative to the onset of the clock is shown in Figure S1.

A sensory temporal estimation task was used to approximate the measurement error in the basic Libet paradigm. The subject was instructed to use the position of the clock dial to report the time of an event using the methods described above. However, in the sensory test, the subject was asked to report the time of a short sound stimulus (250ms auditory beep) that was played between 2.3 and 9.2 seconds (randomly sampled from a uniform distribution) after the clock presentation. Other aspects of presentation, such as the revolution rate of the clock, random initial starting position, etc., were identical to the baseline Libet task. A total of 2 sessions in NS and one session in EGS were acquired.

Opt-out paradigm

The Opt-Out task was designed to understand how the subject's high-level intention to participate in the experiment modulated neural activity observed during task performance. The Opt-Out paradigm is similar to the basic Libet Paradigm described above. However, in this variant, we instructed the subject to choose whether to execute the trial or "opt-out" on a trial-to-trial basis. The subject passively observed the clock as usual during opt-out trials until the trial timed out after 16 seconds. The timeout period was chosen based on the participant's choice behavior in the Libet task (Figure S2). We increased the inter-trial interval to five seconds during the opt-out task and asked the subject to decide to participate or not before the effector-cue screen. We asked the subject to report if they could not make up their mind before a beep played 1.5 seconds before the effector cue. In this event, the trial was aborted, and a new trial was initiated. The subject aborted 1 of 286 total trials and confirmed after each session that the decision to perform the trial was made before the effector cue for all non-aborted trials. A minimum of 60 trials were recorded per session. Three sessions were recorded in subject NS. Due to the free-choice nature of the opt-out experiment, one concern is that NS did not follow the instructions. For example, NS may have waited for the effector cue to choose to perform a trial, and thus, we were unable to decode her choice because she had yet to choose. However, NS reported that she consistently chose before a beep played 1.5 seconds before the effector cue. Further, in other studies involving the same participant, we have found strong evidence of behavioral compliance despite an inability to validate behavior externally. For example, we have found classification accuracy approaching 100% in motor imagery tasks, suggesting remarkable trial-to-trial compliance. However, without accompanying evidence showing encoding of the decision to participate in alternate brain regions, and in the absence of external validation of the precise timing of the internal thought processes of the subject, we acknowledge that future work is needed to understand possible encoding of abstract decision variables.

Simple choice paradigm

The Simple Choice Paradigm (SCP) was designed to observe neural behavior in conditions with minimal task structure (Figure 3A). The subject was free to move the shoulder or attempt movement of the thumb or index whenever she felt the urge to do so. The participant was instructed to say "shrug," "thumb," or "point" timed to movement onset to indicate the timing of otherwise unobservable movements (e.g., thumb and index movements are below the level of injury and thus are not overtly produced.) Movements could thus be considered the conjunction of two actions, executed or attempted movement of the hand and arm, and a vocalization.

Movement onset time was determined by the onset of an audio trace recorded by the NSP, synchronized to neural activity. In addition, we recorded shoulder EMG activity to verify the relative onset timing of vocalization and associated movement. Detection of the participant's verbal report was displayed as a brief (250 ms) change in the color of a simple annulus shown on the screen. The subject was instructed to move as soon as they experienced an urge to move. However, they were asked to avoid repeated sequences of movements such that the movements were "spontaneous and unpredictable." The subject did not report the time they experienced this urge to allow free-flowing progression of movements without interruption. A minimum of 180 trials were recorded per session. Five sessions were recorded from subject NS.

Please cite this article in press as: Aflalo et al., Implicit mechanisms of intention, Current Biology (2022), https://doi.org/10.1016/j.cub.2022.03.047

Analysis methods

Principal component analysis

Principal component analysis (PCA) was used to summarize the population-level temporal response of simultaneously recorded neurons.[66] We constructed a matrix of neural data D that was (n) by (t * c) in size, with n being the number of neurons, t being the number of time points, and c being the number of conditions. For each neuron, activity was averaged across repetitions of the same condition, smoothed with a 250ms box-car function, sampled every 10 ms, normalized to a range between 0 and 1, and mean subtracted. Principal components were calculated based on the singular value decomposition algorithm resulting in $\widehat{D}$ =T*D. To display single-trial representations or the standard error of the mean for all trials of a principal component (e.g., Figure 1B), we applied the transformation matrix T (computed based on trial-averaged data) to single-trial data preprocessed in the same manner described above, though with normalization terms set by the trial-averaged data. The details of how D was constructed depended on the dataset. For the analysis related to Figure 1B, we used a single temporal window extending from -3 to 2 seconds relative to the time of conscious intent (W) and a single condition (shoulder shrug trials) similar to previous studies e.g.[5–7,9,19] We focused analysis on shoulder trials because movement prior to W could be detected by simultaneously acquired EMG measurements. Only one trial was found to have early EMG activity (<1%), and further, the participant self-reported that she had moved early.

Cluster analysis of neural responses

Single neurons varied in their temporal responses (e.g., Figure 1C) We applied a cluster analysis to determine whether the large number of recorded neurons could be categorized into a smaller set of basic temporal profiles. We limited the analysis to neurons exhibiting significant temporal modulation through time. Neural activity extending from 2 seconds before to 2 seconds after W was binned into 50 ms windows. The average response was computed across trials, and the result was normalized to a range from 0 to 1. We constructed a matrix from the resulting temporal responses that were n x t in size, with n being the total number of neurons and t being the total number of time bins. We applied Gaussian mixture model cluster analysis to the matrix using k=1 through 10 possible clusters, each with 500 randomized initializations of cluster centers. For each k, we computed the Bayesian information criteria (BIC) to determine the optimal number of clusters. Different cluster identification methods can result in different numbers of clusters being identified as "optimal." However, the objective of this analysis was not to determine the "true" number of clusters but instead, to use a principled method to group neurons with similar temporal profiles. This enabled testing of whether the timing of a neuron's response could explain whether a unit was better associated with conscious intent or motor production.

Timing analysis of neural signals with respect to M and W

The objective of this analysis was to determine whether neural signals are better aligned to the time the participant reported being aware of the urge to initiate movement (W) or the time of movement production, defined as muscle activity onset. We computed the average temporal profile for each neuron over a three-second interval, binned into 250 ms windows, centered on M or W. We then computed the difference in trial-to-trial variance explained by the mean profile. Significance was calculated by comparing this difference to an empirically computed null distribution of differences using a rank test with alpha = 0.05. The null distribution was generated by repeatedly (1000 permutations) calculating the difference in variance explained between two datasets generated by randomly assigning M-aligned and W-aligned trials. The mean offset between M and W trials was removed before random assignment. A schematic representation of this analysis is illustrated in Figure S2. This analysis was performed on each unit separately, and the results are reported categorized by the clusters identified in section cluster analysis of neural responses (see Figure 1D). Potential complications in interpreting this analysis include inherent measurement error differences between M and W. See Figure S3 for relevant results and discussion.

Basic population analysis

We used a linear model to quantify the percent of the population tuned to the salient task variables in the basic Libet paradigm. Neural activity was divided into three temporal epochs centered on the time of cue presentation, clock presentation, and the reported urge to initiate movement. Neural activity was averaged in 500 ms windows at 100 ms intervals (sliding window analysis.) Each window was compared to baseline activity during the inter-trial interval, chosen as a window -1000 to 0 seconds before cue onset. For each time bin, we used linear regression to model neural firing rate relative to baseline as a function of the response to the two effectors. We then looked at whether either effector changed significantly from baseline and at the contrast comparing the activity of the two effectors (to find units demonstrating effector specific modulation.) Percent significant (e.g., Figures 1F and 2C) are shown uncorrected at $p<0.05$. We performed a permutation test to determine whether the percent of units across the population was significant by using a rank test to compare the uncorrected percent significant against an empirical null distribution of percent significant generated by shuffling labels and repeating analyses 2000 times.

Cross-validated demixed principal component analysis

We used demixed principal component analysis (dPCA) to visualize effector-specific and independent effector components of the neural response.[27,67] Demixed principal component analysis is a supervised dimensionality reduction technique that uses information about task parameters to find the low-dimensional latent factors that best separate task parameters. Similar to classification algorithms, dPCA has the potential to overfit training data, finding low-dimensional projections of the neural data that separate task labels in a way that won't generalize to data that is not used to compute the dPCA projection matrices.[27] We, therefore, used a stratified hold-one-out cross-validation scheme to verify the validity of the projections. We applied regularized dPCA[27] to compute dPCA projection matrices on all but one trial from each condition. The resulting projection matrices were applied to the held out trials to compute the generalized latent response estimate. This process was repeated for all combinations of held-out trials, and the mean and 95% confidence intervals were estimated using a bootstrap procedure. Trial data was constructed by binning neural firing

into 500 ms non-overlapping windows and concatenating in time data from the cue aligned, clock aligned, and W aligned epochs. Analysis was repeated both on trials acquired within a single session and trials pooled across days. For pooled analysis, neurons acquired in separate sessions were treated as independent. Results in Figure 1G are shown for pooled data. Single session data were qualitatively similar.

Basic classification methods

Our base approach to classification was to use linear discriminate analysis. To regularize covariance estimates and improve cross-validated accuracy, we modeled the covariance of the normal distribution as diagonal and pooled our data to create a single covariance estimate that was applied to all conditions. These simplifying assumptions were found to improve cross-validation prediction accuracy on preliminary data. The classifier input was a matrix of average firing rates within a specified window for each sorted unit. Classification performance is reported as generalization accuracy of a stratified leave-one-out cross-validation analysis. Features that demonstrated non-significant tuning based on a preliminary ANOVA test were excluded from the input vector to reduce the total number of features. For cross-validation purposes, the ANOVA test exclusion criteria were calculated on the training set and applied to the test set to avoid ''peaking'' effects. We used a shuffle procedure to determine the significance of decode accuracy. A null distribution of decoding accuracies was generated by shuffling the task labels associated with each feature vector and repeating the cross-validated decode analysis. This process was repeated 500 times, and significance was determined by a one-sided rank test, testing whether the veridical accuracy of the unshuffled data was greater than 95% of the shuffled results. Analysis was repeated both on trials acquired within a single session and trials pooled across days. For pooled analysis, neurons acquired in separate sessions were treated as independent. Results were qualitatively similar, and thus, we used pooled data for reporting purposes.

In preliminary analysis, we used alternate classification methods. This included relaxing some of the assumptions of linear discriminant analysis as performed above: allowing off-diagonal elements to the covariance matrices and allowing separate covariance estimates for each class. We also tried naïve Bayesian classification and linear, quadratic, and Gaussian support vector machines with cross-validated optimization of parameters: we would find the set of parameters that optimized cross-validated performance across all but one and test the parameters on the cross-validated performance on the remaining session. Linear discriminant analysis, as described above, outperformed all other tested methods when classification was performed on a predetermined window of time. However, when the precise window of time was unknown (e.g., when the classifier was applied continuously at each time step as is necessary for closed-loop performance), alternative techniques were required to improve performance as described below.

Decoding choice to participate

In the opt out paradigm, the subject chose to perform any given trial before the effector cue. Is information about the high-level intention to participate in a given trial encoded in the neural population? Each trial was labeled as either a ''Go'' or ''NoGo'' trial based on whether the subject chose to perform the trial. We then used a classification analysis during four epochs to determine whether and when information related to the subject's intent to participate was encoded in the neural population. These epochs include a ''Pre Cue'' epoch (mean response between -750 and 0 ms relative to the effector cue), a ''Post Cue'' epoch (mean response between 250 and 1000 ms relative to the effector cue), a ''Clock'' epoch (mean response between 250 and 1000 ms relative to clock onset), and a ''Go'' epoch (mean response between -250 and 500 ms relative to reported urge to initiate movement). For NoGo trials, activity for the ''Go'' epoch was generated by averaging data in a 750 ms window chosen an interval of time after clock onset based on movement production times of ''Go'' trials (e.g., Figure S2). We used cross-validated LDA to decode Go from NoGo trials for each epoch. After the effector cue, the neural population encodes a motor plan (as evidenced by effector specificity). Thus, Go versus NoGo trials can be distinguished based on the presence or absence of the motor plan. For this reason, the most parsimonious explanation of the ability to decode Go from NoGo trials after the effector cue is simply the presence or absence of the motor plan. Either way, the ability to differentiate Go from NoGo trials (Figure 2D) using the techniques described above validates the decoding methodology.

Mahalanobis distance and classification through time

Classification using LDA is based on relative distances in Mahalanobis space.[68] Each class is defined by a mean and covariance in feature space. To classify a point in feature space, the distance between the point and the mean of each class normalized by the covariance matrix is computed. The point is assigned to the class that results in the shortest distance. More precisely:

$$\hat{y} = argmin(y = 1, \ldots K) \sum_{k=1}^{K} P(k|x)$$

Where $\hat{y}$ is the predicted class, K is the number of classes, x represents the features, $P(k|x)$ is the posterior probability of class k for observation x and is given as:

$$P(k|x) = \frac{P(x|k)P(k)}{P(x)}$$

and

$$P(x|k) = \frac{1}{\sqrt{2\pi|\Sigma_k|}} \exp\left(-\frac{1}{2}D_k^2\right)$$

$$D_k = \sqrt{(x - u_k)^T \sum_k^{-1} (x - u_k)}$$

D_k defines the Mahalanobis distance from the data point x to the class k defined by the mean and covariance (u_k, Σ_k) and determines the output of the classification procedure (save for normalization terms.) Viewed in this way, LDA simply discretizes the Mahalanobis distance, assigning the data point to the class with the smallest Mahalanobis distance.

We computed a relative distance measure as a normalized difference of Mahalanobis distances to better characterize how population activity behaved through a trial.

$$Relative\ Distance = \frac{D_{ITI} - D_{GO}}{D_{ITI} + D_{GO}}$$

Where D_{ITI}&D_{GO} define the distance from a data point to the classes that characterize the intertrial interval and execution-related activity (Figure 4A). The sign of $D_{ITI} - D_{GO}$ determines whether a data point would be classified as ITI or GO activity. Note that with trial-to-trial variability, the closer the average difference is to zero, the more likely responses on a given trial will be misclassified. Relative distances are normalized into the range of -1 to 1 by dividing by $D_{ITI} + D_{GO}$. By looking at a normalized version of Mahalanobis distance, we can better understand what drives classifier behavior. For instance, task events might make it more likely to decode class B when the true class is A either because the task events lead to greater variability in the neural features (but leave the mean response of the neural features the same) or because the task events drive the mean response of the neural features closer to class B (but leave the variability largely similar) or some combination. Looking at neural behavior in Mahalanobis space can thus lead to greater understanding and interpretability of decoder output. A schematic representation of this analysis procedure is illustrated in Figures 4A and 4B.

Mean and covariance estimates for the ITI period were computed based on neural features averaged within a 1-second window -1.5 to -0.5 seconds before the effector cue. Mean and covariance estimates for the Go period were calculated separately for shoulder shrug and hand squeeze trials based on neural features averaged within a 1-second window 0 to 1 second relative to W. Data points throughout the trial were constructed by computing the average firing rate within 500 ms non-overlapping windows and concatenating in time data from the cue aligned, clock aligned, and W aligned epochs for both trials where the subject was cued to perform a shrug movement and squeeze movement. The relative distance measure was then computed for each such data point. There were clear phasic responses following the cue and around the time of movement production. Does the activity in between these phasic responses change? We compared the relative distance measure for the windows starting 1 sec after cue onset and 1.5 seconds before the reported urge to move using permutation rank test where the magnitude of the difference computed for the two windows was compared to a null distribution of means generated by randomly shuffling data between the two time windows.

The same data was used for linear discriminant classification analysis by transforming Mahalanobis distances into class estimates as outlined above. As described above, classification accuracy was estimated using a stratified leave-one-out cross-validation procedure. In this case, the entire time-series comprising a trial was held out to ensure that autocorrelation in the signal did not provide information about adjacent time slices. Each time slice for each trial was classified as shoulder, squeeze, or ITI. The proportion of trials assigned to each class was displayed as the relative height of a colored bar, the heights of which necessarily sum to 100% (e.g., all trials).

Improving classification

Population decoding has become a ubiquitous technique to understand how populations of neurons encode sensory, motor, and cognitive variables. In contrast to common uses of classification algorithms for offline data analysis, online decoding algorithms must determine both what the participant intends and when the participant intends action. As shown in Figure 4, certain neural processes, such as encoding of the motor plan in response to a cue, are sufficiently similar to neural activity patterns recorded during motor execution to generate early spurious decodes. Here we show that classification performance is improved by leveraging information from early encoding and maintenance of motor plans combined with classification algorithms that leverage temporal history (e.g., compare Figures 4D and 4E).

To motivate the updated approach to neural classification, we present the underlying issues in a toy dataset in the supporting information (Figure S4). The dataset is composed of the same neural recordings used for the analysis shown in Figure 4 of the main text; however, for illustrative purposes, visualization and decoding are based on the first two principal components of the full-dimensional neural signals. Further, the activity of each ''trial'' is constructed by averaging the activity of two repetitions of the condition. This aids visualization and shows that the underlying problem is not the product of noisy data (although noisy data exacerbates the underlying issues). The results for all neural dimensions, unaveraged, are shown in Figure 4 with visualization of the response in panel c and decode performance in panels d and e.

In standard analysis-based approaches to neural classification, several conditions are tested, and the associated neural population response is sampled in a predefined temporal window to generate features for classification. A classification algorithm then partitions the neural feature space into separate regions defined by decision boundaries. An example illustrating this process for three conditions is shown in Figure S4A. One condition represents the ''null-state'', the activity of the neural population sampled during the inter-trial interval when the subject intends no action (gray). The other two classes represent the ''go'' period activity measured during the execution of two movements (movement 1/M1 = green, movement 2/M2 = red). The decision space that separates these classes

was computed using linear discriminant analysis, and each condition is color-coded. Leave-one-out classification of data acquired from these predetermined windows is perfect. However, if neural data just following the cue is projected onto the same decision space, a significant portion of trials have crossed the decision boundary and thus would be decoded as a movement, even though the subject does not intend to initiate a movement (Figure S5B). Spurious decoding when a simple LDA classifier trained on execution-related neural responses is pervasive when the decoder is applied continuously throughout the trial (Figures S4C and S4D; see also Figure 4D). The two panels show the classification results for movement one trials (top) and movement two trials (bottom), with each bar showing the leave-one out classifier output (movement 1, movement 2, or null) for each slice of time (500ms non-overlapping windows as in Figure 4). There is evidence of premature decoding or the correct movement and spurious decoding of the alternate movement. The behavior of the neural population generating the classifier output can be visualized throughout the temporal interval by animating the continuous neural dynamics on the decision space (Video S3). The problems are twofold. First, planning and decision-making activity early in the trial generates changes in neural activity that can cause the neural population to enter unwanted states. Second, execution-related neural activity can trace a trajectory through time that brings population activity generated by one movement close to the decision boundary defined by another movement.[69–71] Including more neural dimensions than the two used in the toy, dataset can help classification accuracy. Nonetheless, even when including all relevant neural dimensions, neural dynamics can lead to problematic classifier output (e.g., Figure 4).

We took two steps to fix the identified problems algorithmically. First, we define additional neural states in time that can capture neural activity patterns not directly related to the intention to initiate movement (e.g., sensory, decision making, and planning activity.) This better partitions the neural state-space which leads to a more precise understanding of how the neural state relates to the intention to initiate action. The second step was to replace a basic linear discriminant classifier with Hidden Markov Model (HMM) decoding.[68,72] Essentially, a HMM (as employed here) has two components. The first is an instantaneous estimate of the class based on direct classification of neural data using the same procedure as LDA as described above. The second component leverages temporal history in the neural signals in the form of the recent state history (modeled as a Markov process).

The state transition model used for decoding is shown in Figure S4E. Initial state transition probabilities were fit from the data. The probabilities reflect the task design (e.g., the length of task epochs and transitions between these epochs) and the choice of window size (e.g., 500ms non-overlapping windows). The effect of the HMM with multiple states applied to the data from Figures S4A–S4D is shown in Figures S4F–S4I and Video S3 (inset). In particular, the finer partitioning of the neural state space is evident in panels f and g. The net impact on cross-validated classifier output is shown in panels h and i.

Early intent decoding in simple choice paradigm

The simple choice paradigm was used to address a simple question: What is the timing of motor planning signals when subjects are free to choose when and what to move without any strictly imposed task structure? To address this question, we aligned neural data to movement onset. We used classification analyses to decode both the movement the participant is actively executing and the next movement the subject will execute in a sliding window analysis. Windowed data was computed at interval steps of 50 ms using the average firing rate of each neuron within a 500 ms window centered on the time point of interest (sliding window analyses.) For each trial, two sets of labels were generated: one label associated with the current action the subject performed (current movement) and a second label associated with the movement the subject reported performing on the next trial (next movement). Two independent cross-validated decoding analyses were performed, one for each set of labeled data. We used LDA classification with constraints as described above.

A shuffle test determined significant decoding. We first calculated a veridical estimate of accuracy using the task labels. We next computed a null distribution by shuffling task labels and recomputing an accuracy estimate 500 times. A single time slice was determined to be significant if the veridical estimate was greater than 97.5% of shuffled trials (rank test, $p<0.025$.) To determine the time at which the population first encoded the next chosen movement, we looked for the time slice that started a continuous block of significant decoding: this block was required to be 15 time slices in duration (750 ms) to ensure that non-overlapping bins of neural data demonstrated significant decoding. We circularly shifted the neural data relative to behavior randomly, with a minimal shift of 28 seconds, and repeated the analysis 100 times to verify this procedure. We did not find a significant block of 15 time bins in any of the 100 repetitions. We know that neural activity can dramatically rise >1s before movement initiation in PPC (e.g., Figures 1B and 1C). To ensure that this sudden rise cannot explain a significant early rise in activity starting at the previous trial, we split the trials into long and short interval trials and performed the analysis separately.

Statistics

No statistical methods were used to predetermine sample size. The experiments were not randomized. Data analyses were performed by automatic software routines. All data that was recorded was included in the results. No sessions or units were removed from the study based on analysis results, behavior, or other criteria. Non-parametric permutation testing was used throughout to avoid the assumption of normality.