HHS Public Access

Author manuscript

Nat Neurosci. Author manuscript; available in PMC 2013 June 01.

Published in final edited form as:

Nat Neurosci. 2012 December ; 15(12): 1715–1722. doi:10.1038/nn.3250.

Neural population partitioning and a concurrent brain-machine interface for sequential motor function

Maryam M. Shanechi[1,2,3], **Rollin C. Hu**[2,3], **Marissa Powers**[2], **Gregory W. Wornell**[1], **Emery N. Brown**[3,4], and **Ziv M. Williams**[2,3]

[1]Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge, MA, USA

[2]Department of Neurosurgery, Massachusetts General Hospital, Boston, MA, USA

[3]Harvard Medical School, Boston, MA, USA

[4]Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, Cambridge, MA, USA

Abstract

While brain-machine interfaces (BMIs) have largely focused on performing single-targeted movements, many natural tasks involve planning a complete sequence of such movements before execution. For these tasks, a BMI that can concurrently decode the full planned sequence prior to its execution may also consider the higher-level goal of the task to reformulate and perform it more effectively. Here, we show that concurrent BMI decoding is possible. Using population-wide modeling, we discover two distinct subpopulations of neurons in the rhesus monkey premotor cortex that allow two planned targets of a sequential movement to be simultaneously held in working memory without degradation. Such surprising stability occurred because each subpopulation encoded either only currently held or only newly added target information irrespective of the exact sequence. Based on these findings, we develop a BMI that concurrently decodes a full motor sequence in advance of movement and then can accurately execute it as desired.

An important motivation for the design of brain-machine interfaces (BMIs) to date has been their potential ability to restore lost motor function in individuals with neurological injury or disease (e.g., due to motor paralysis or stroke). In such cases, the envisioned role of the BMI is to decode the intended movement from neural activity in the relevant areas of the brain, and use this information to control an affected limb, prosthetic, or other device.

Correspondence should be addressed to Z.M.W. (zwilliams@partners.org). .

Note: Supplementary information is available on the Nature Neuroscience website.

AUTHOR CONTRIBUTIONS M.M.S. developed the real-time decoder, conceived and performed the computational analysis, assisted with animal recordings, and wrote the manuscript. R.C.H. and M.P. assisted with animal training and recordings. G.W.W. and E.N.B. were involved in the computational methodological development and edited the manuscript. Z.M.W. conceived and designed the study, developed the BMI system, performed the animal training and recordings and wrote the manuscript.

COMPETING FINANCIAL INTERESTS The authors declare no competing financial interests.

The design of such BMIs has received considerable attention in recent years[1-18]. Work to date has principally focused on achieving the motor goal in tasks that involve single-targeted movements, such as the task of moving a cursor on a display to an individual target location. These BMIs can decode the continuous trajectory of one- to three-dimensional movement (including a grasp in some studies)[1-14], the intended target location[15, 16], or both the target and trajectory jointly using approaches such as optimal feedback control[17, 18]. However, in many natural tasks—such as playing a succession of notes on a piano—the goal is more complex, and the motor plan for achieving it can be viewed as a complete sequence of such simpler plan elements to be executed in order.

Our focus is on the design of BMIs that can achieve the goal of these sequential motor plans. Planned sequential behavior is a fundamental motor process in which all targets of a movement sequence are planned ahead of its initiation. Hence a BMI for performing such behavior would allow a person to *plan a full motor sequence* ahead of execution. For example, when picking up a cup and bringing it to one's lips, a person normally formulates the complete motor plan prior to its execution as opposed to planning and performing each of its elements individually and separately. Therefore, the objective of such a BMI would be to perform the sequential behavior by decoding all elements of the sequence concurrently and in advance of movement – thus requiring the consideration of a *concurrent* architecture. This BMI functionality is distinct from that in prior BMIs that decode and execute individual single-targeted movements one by one, and hence have a *sequential* BMI architecture[1-18].

In addition to simultaneously decoding a motor sequence in advance, a concurrent architecture could also allow the BMI to consider the overall motor goal of the task at a higher-level. This is a result of the BMI having information about all the motor plan elements at once and in advance of execution. Hence one prospective BMI capability would be to consider all elements of the sequence concurrently, prior to action, in order to determine ways to perform the task more effectively. For example, the BMI might determine a way to accomplish the task more quickly, or more efficiently (to within any physical constraints that might exist). Alternatively, based on additional sensor inputs, the BMI might determine that the planned sequence of movements would result in an accident with an obstacle, and thus modify the execution of the task to avoid such an accident.

The realization of BMIs that can perform and potentially execute sequential motor function more effectively in this way will obviously require significant technological innovations. But as a key initial step, it requires considering a concurrent BMI architecture in which the elements of a planned motor task are decoded in parallel (i.e., at once), in contrast to the serial process of a sequential BMI. Hence, the feasibility of such BMIs hinges on the degree to which the elements of a motor plan sequence can, in fact, be decoded concurrently. This is the starting point for our research.

Prior work has demonstrated that individual neurons in the premotor cortex of primates display selective responses to planned single-targeted movements before their initiation, and that such responses often remain sustained during working memory until movement execution[19-26]. Such responses have been successfully exploited in the design of BMIs for

single-target tasks[15, 16]. In comparison, the neural encoding of tasks requiring a full sequence of planned targeted movements to be formulated prior to execution is less well understood, and the design of real-time BMIs that can concurrently decode and then execute such sequential motor plans remains unexplored. Prior work has shown that an individual neuron can display a response that is selective to one or more elements of a sequential motor plan[27-41] (see also Discussion). However, little is known regarding how information about multiple elements of a sequential motor plan (e.g., the planned targets of a sequential movement) is simultaneously distributed across the whole premotor population during working memory, and whether these plan elements can be accurately decoded from the neural population in a concurrent manner. More importantly, it is necessary to determine whether adding information about the elements of the motor plan, in sequence, to working memory affects the integrity of information about the plan elements that are already held, and how it affects their neural encoding. Finally, it is necessary to assess robustness—whether a BMI limited to recording from relatively small numbers of neurons is able to achieve sufficient and consistent decoding accuracy.

Here, we find that sequential motor plans can be decoded simultaneously, accurately, robustly, and in advance of movement from the neural activity in the premotor cortex of monkeys. Additionally our study reveals a surprisingly structured encoding mechanism that is used by the premotor populations for these sequential plans and that, in turn, allows for their accurate and concurrent decoding. Based on these findings, we develop and implement a real-time BMI that can concurrently decode a dual sequence of motor targets and then execute them as desired.

RESULTS

In the study, two adult male rhesus monkeys were trained to perform a task in which two targets were presented, in sequence, on a computer display. Each of the targets could randomly take on one of four possible spatial locations (“up”, “down”, “left”, or “right”). Repeated locations were precluded, so there were a total of 12 possible combinations (sequences) of two consecutive distinct target locations. After a blank-screen variable delay, a “go” cue appeared directing the monkeys to sequentially move a cursor from the center of the screen to each of the two remembered targets, in order (Dual-target task; Fig. 1a, b). We define the working memory period as the 500 ms blank-screen interval following presentation of the second target and before the earliest possible “go” cue. Therefore, the task here was a working-memory task in which the monkeys were required to serially add to working memory two *randomly* selected target locations in each trial and then simultaneously retain them in working memory prior to execution.

Multiple-unit responses were recorded from the premotor cortex as the primates performed this task. We recorded 281 well-isolated single neurons from the supplementary motor area (SMA) and dorsal premotor cortex (PMd) over 11 sessions, for an average of 26 ± 6 cells (mean ± s.d.) per recording session (note that some of these cells may not be distinct across the different sessions). Inhomogeneous Poisson models were fitted to each neuron’s spiking activity using an expectation-maximization algorithm[42] (see Methods and Supplementary Modeling). Using these models, we employed a maximum-likelihood decoder to quantify

the probabilities that groups of neurons could correctly identify the first and second targets on a trial-by-trial basis during the working memory period (leave-one-out cross-validation; see Methods). We used decoding accuracy as our measure of the amount of information encoded by a population of neurons about each target. Specifically, for an individual (first or second) target, we measured the percentage of trials in which the maximum-likelihood decoder correctly predicted the respective target from that population's activity. Likewise, we measured the amount of information encoded about the full sequence as the percentage of trials in which both targets were correctly decoded.

Accurate and concurrent encoding of the motor sequence

We find that neural population activity within the premotor cortex accurately encoded the location of both targets during the working memory period. During this period, the population correctly encoded the first and second targets on 85% and 82% of the trials in the best session, respectively. When considering all 12 possible target combinations, the population encoded both targets correctly on 72% of the trials in this session (Fig. 2a; 285 dual-target trials were performed in this session). Across all tested sessions, the population correctly encoded the first and second targets on average on 76 ± 11% and 56 ± 17% of trials, respectively, both of which were significantly above chance (one-sided Z-test, $P < 10^{-15}$; Supplementary Fig. 1a). Also, the population encoded both targets correctly on average on 45 ± 12% of the trials across all sessions, which was also far higher than chance at $1/12 \approx 8\%$ (one-sided Z-test, $P < 10^{-15}$). These results were consistent across the two monkeys ($P < 10^{-15}$ for both; Supplementary Fig. 2).

Robustness of the encoding

Only a small number of cells were needed to decode the target sequence with high accuracy. When performing the decoding analysis over all trials, which employed all 12 possible target combinations, only 29% of the population (7.5 cells) was needed, on average, to achieve higher than 90% of the population sequence accuracy (Fig. 2b; see Methods). When performing the decoding analysis over subsets of all trials that employed only 4 or 8 target combinations, population sequence accuracies in the best session were as high as 93% and 80%, respectively. In these cases, decoding from only 2 and 4 cells, respectively, was sufficient to achieve higher than 90% of these sequence accuracies.

Real-time concurrent BMI for sequential movement execution

Motivated by the observation that both targets can be concurrently and accurately decoded from the responses of relatively few neurons in the premotor cortex, we developed a real-time BMI capable of predicting both targets simultaneously prior to monkey's motor response and then executing the targeted movements. In the associated experiments, we recorded a mean of 20 ± 2 cells per session from the premotor cortex of the same monkeys. Here Poisson models were first fitted to the neural population activity during the working memory period prior to the "go" cue (see Methods and Supplementary Fig. 1b) as the primates rehearsed a subset of target combinations that included either 4 or 8 possible sequences over an average of 26 ± 2 training trials per sequence (Fig. 1b). We chose to use

either four or eight sequences in the BMI experiments to obtain sufficient training and real-time trials per session.

Using the Poisson models, sequence decoding accuracies for the set of 4 and 8 sequences in these training sessions (found using leave-one-out cross-validation) were 79 ± 2% and 80 ± 3% (mean ± s.e.m.; one-sided *Z*-test, $P < 10^{-15}$), respectively. Following training, the primates performed the same task as before, but with the cursor now being sequentially positioned by the BMI on the targets decoded from the recorded neuronal activity during the single preceding working memory period (Fig. 1c; see Methods). Here, BMI-generated cursor movements were set to occur immediately following the presentation of the "go" cue and the added delays were selected to match the reaction times that the monkeys normally experienced when moving the cursor themselves following the "go" cue (obviously, cursor movements could be generated without the added delays if desired).

Both monkeys performed a total of 459 trials on the four-sequence set, and one monkey performed 110 trials on the eight-sequence set using the real-time BMI. Sequence accuracies for the set of four and eight sequences were 72 ± 2% and 71 ± 4%, respectively, both of which were significantly above chance (mean ± s.e.m., one-sided *Z*-test, $P < 10^{-15}$). Both training and real-time BMI accuracies were similar and significantly above chance across the two monkeys (one-sided *Z*-test, $P < 10^{-15}$ for both; For the four sequence sets, the first monkey had a BMI accuracy of 69 ± 3% and a training session accuracy of 77 ± 2% and the second monkey had a BMI accuracy of 75 ± 3% and a training session accuracy of 82 ± 2%). Sequence accuracies using the BMI were also close to the cross-validated sequence accuracies during the training sessions when taking into account the primates' natural error rates during the standard task (Fig. 3). In fact the 95% confidence bounds for the two accuracies were overlapping (72 ± 4% vs. 73 ± 3% and 71 ± 8% vs. 66 ± 6%, for the sets of four and eight sequences, respectively; see Methods). These results established that two planned elements, i.e., the two intended sequential targets of movement, could be simultaneously predicted in advance of movement and then executed by a real-time BMI with high accuracy.

We also examined the time required by the concurrent decoder to decode the sequence. We find that the sequence decoding accuracy for the set of four, eight, and twelve sequences reached 90% of the maximum asymptotic accuracy possible, on average, after 488 ± 135 ms, 561 ± 119 ms, and 641 ± 121 ms from the *initial* presentation of the second target, respectively (Fig. 4). When performing the motor sequence, the minimum total time it took for the monkeys to both react to the two "go" cues and reach the two targets was on average 791 ± 93 ms (this is the sum of the *two* reaction times *plus* the *two* center-to-target movement times; Fig. 4).

Population encoding reveals a novel partitioning mechanism

Observing that both target locations could be accurately and concurrently predicted from the neural population responses, we further examined the spatial and temporal structure of their encoding. In particular, we investigated how neurons within the premotor cortex were able to add new information about the second target to working memory without compromising the integrity of information about the first target that was already being held. To do so, we

used the decoding approach that measures the amount of information held about the identity of each planned target in the sequence by considering *all sequence combinations collectively*.

We find that most cells encoded significant information about *only* the first (currently held) or *only* the second (newly added) target during the working memory period. Moreover, this partitioning was present across all target locations/sequences (i.e., responses were not sequence specific) and remained stable throughout recordings per day. Of the 281 neurons recorded in all sessions, 46% had a target accuracy significantly higher than chance for at least one of the two targets during the working memory period (one-sided *Z*-test, $P < 0.01$). Of these, 68% encoded significant information about only the first currently held target (Supplementary Fig. 3), and 23% encoded significant information about only the second added target (one-sided *Z*-test; $P < 0.01$; Fig. 5). The percentage of cells that encoded significant information about both targets was only 9% (one-sided *Z*-test; $P < 0.01$; note that a Bonferroni correction for multiple comparisons was done for *all* comparisons; Supplementary Fig. 4) and, even among these, most had target accuracies much closer to one of the two subpopulations of cells that significantly encoded only one target (Fig. 6; Supplementary Fig. 5). These results revealed a highly significant divergence in the amount of information encoded by the two subpopulations of neurons about the two targets (random permutation test, $P < 10^{-15}$; see Supplementary Modeling; Supplementary Fig. 6). Moreover, we examined the relation between the activity of each of the two subpopulations to upcoming motor behavior and found that each subpopulation was only predictive of whether the respective first or second upcoming movement would be performed correctly or incorrectly (i.e., resulting in a behavioral error) by the primates following the “go” cue (one-sided *Z*-test, $P < 10^{-15}$).

These results demonstrate that during the *working memory* period, most neurons were not selective to a specific sequence or simply to a spatial location. Rather, they were partitioned into two disjoint subpopulation, one encoding only the identity of the currently held (first) target and one encoding only the identity of the newly added (second) target within the sequence, *regardless of the specific sequence* (see Supplementary Fig. 7 for comparison to sequence specific selectivity found in prior work[28-31, 41, 43]).

The observed partitioning during working memory was not related to limb movement or simple visual related responses. No visual cues were presented during the working memory period and any movement before the “go” cue terminated the trial. This was also suggested by the partitioning mechanism itself since if the activity was due to targeted limb movement, then all cells would only reflect the direction of this single target. Finally the electromyography (EMG) activity during the working memory period was not predictive of the first movement direction (one-sided *Z*-test, $P = 0.14$), but was predictive of it during the first movement period after the “go” cue (one-sided *Z*-test, $P = 0.01$). In an additional set of analyses, we also found that encoding of the second target was not conditioned on the location of the first target, and vice versa (Supplementary Fig. 8).

Effect of adding information to working memory

In order to further examine how adding a new target to working memory affected the integrity of the currently held target, we disambiguated the process of holding information in working memory from that of adding information to it. The results were obtained from sessions in which the monkeys performed the standard dual-target trials (as before), but also performed single-target trials in randomly interleaved fashion (see Methods). Unlike dual-target trials, on single-target trials only the first target was presented and the second target presentation period was replaced with a blank-screen period of the same duration. The task timing was otherwise unchanged compared to the dual-target task.

We find that adding information about the second target location to working memory did not incur loss of information about the first target location. Of the cells that encoded significant information about the first target during working memory in single-target trials (one-sided *Z*-test, $P < 0.01$), most (74%) provided the same level of accuracy in decoding the first target during working memory in dual-target trials, despite the addition of a second target (χ^2 test, $P > 0.05$). Moreover, for the whole population, there was also no significant difference in the first target accuracy during the working memory period when comparing dual-target and single-target trials across sessions (Wilcoxon's signed-rank test, $P = 0.69$; Fig. 7). These results demonstrate that the subpopulation encoding the first target and their responses remained largely unchanged when the second target was added to working memory and, therefore, the addition of information about the second target did not comprise the integrity of information already held about the first target. It is important to emphasize here that the task involved serially adding *to working memory* two *randomly* selected target locations in each trial and then simultaneously holding them in working memory prior to execution. Such a task is distinct from memory-guided tasks in which the same motor sequence is repeatedly performed from memory after learning, or visually-guided tasks in which movements are serially cued and executed one-by-one[29-31, 41, 43].

In a control analysis, we also examined whether neuronal encoding of the first target was affected by the number of targets presented per trial in a single session (i.e., one target versus two sequentially presented targets) by having one monkey perform only single-target trials. Comparing these single-target only sessions with sessions in which single-target trials were interleaved with dual-target trials on the same day, we found no significant difference between the population decoding accuracies of the first target on single-target trials between the two session types (χ^2 test; $P > 0.15$; Supplementary Fig. 9).

Stability of the neural encoding structure

While implicit in the preceding results, it should be emphasized that as the pair of presented target locations varied over the hundreds of trials typical of a given day's session, most neurons remained dedicated to encoding only the first (currently held) or only the second (added) target. For the two subpopulations of cells that encoded significant information about the respective first and second targets alone, most (89%) provided substantially the same level of accuracy in decoding their respective targets in the first and second halves of the recording session (χ^2 test, $P > 0.05$; sessions included 263 ± 36 dual-target trials on average). Also, the sequence decoding accuracy (across all 12 sequences) of the entire

population did not change over time between the first and second halves of the sessions (Wilcoxon's signed-rank test, $P = 0.37$). Therefore, the partitioned premotor subpopulations appeared to be physiologically dedicated to encoding either the first or the second target added to working memory. Inherently, the neural decoding in our BMI exploited this stability of the two constituent subpopulations to achieve sustainable performance.

DISCUSSION

The purpose of the present study was to examine how multiple planned targets of sequential movement are concurrently encoded as a population by premotor neurons during working memory, and to determine whether the activity recorded simultaneously from multiple single-neurons can be used to concurrently and accurately decode the complete motor plan sequence in advance of movement and in real-time. Three methodological approaches were used to investigate these questions. First, we simultaneously recorded the activity of multiple cells across the whole premotor population. Second, we used an interleaved dual-target/single-target task in order to dissociate the dynamic process of maintaining target-related information in working memory from that of adding new information to it. Finally, we employed a maximum-likelihood decoding approach that allowed us to (i) define an accuracy measure for the amount of information that is concurrently encoded about planned motor sequences and, (ii) examine the spatiotemporal distribution of information across the whole population.

A neural partitioning mechanism

Our results reveal a novel functional structure within the premotor cortex that allowed for accurate and concurrent decoding of two planned motor targets across multiple spatial locations. We find that during working memory, premotor populations are largely partitioned into two fundamentally disjoint subpopulations of cells – one dedicated to encoding only the currently held (first) target and one dedicated to encoding only the newly added (second) target, irrespective of the specific sequence. Moreover, while the two target locations changed from trial to trial, the two encoding subpopulations did not. Notably, the subpopulation dedicated to encoding the first target and their responses remained largely unchanged when the second target was added to working memory, so that the process of adding information did not compromise the integrity of existing information (across all target locations). Also, only a small number of neurons were sufficient to accurately predict the location of both targets, making the decoding of such information highly robust.

Prior work has shown that individual premotor neurons display selective responses to single-targeted movements before their initiation[19-26]. It has also been shown that PMd neurons can be selective to the location of multiple target choices for a single-targeted movement before a final selection is made[44], or can represent combined information about the target and the body-part to be used for a single-targeted movement[45-47]. When performing a planned sequential movement, prior studies have demonstrated that individual neurons within areas such as the parietal, premotor and prefrontal cortex can display selective responses to a sequential motor plan[27-40, 43, 48]. Some neurons (often a relatively small fraction) display increased activity for a specific combination of movements (for example a

push followed by a pull of a manipulandum) during a preceding delay, suggesting that they encode information about more than one motor plan element at a time[28-31, 37, 41, 43]. Other cells have also been found to display selective responses during movement itself with increased activity prior to performing a particular movement (e.g., a push) only when it follows another specific movement (e.g., a pull) in sequence[29, 31, 43], or prior to a movement only if it has a particular order in the sequence[31, 43]. What has remained unclear, however, is how information about individual elements of such sequential plans is *simultaneously* distributed across the whole population during *working memory*, and whether and how the process of adding *new information* about an element to working memory affects the integrity of information already held and its neural encoding.

While a major focus of this study was to investigate the encoding structure of premotor populations during the working memory period, we find that, consistent with prior studies[29, 31, 43], neurons often altered the degree to which they encoded information about the two targets across different time-points during the task. Some cells, for example, encoded no information about the second target during the working memory period, but then encoded significant information about the second target during the second movement itself (Supplementary Fig. 5a). Such shifts in activity may reflect the dynamic role premotor neurons play in processing, maintaining, and then executing motor plans in combination with other motor cortical areas.

Another question arising from the study is how information encoded by premotor neurons is related to the later execution of the sequential task. Here, we find that the subpopulation of cells that predominantly encoded information about the first target was only predictive of whether the primates would perform the first upcoming movement correctly or incorrectly, and this was similarly true for the second subpopulation. This thus suggests that the 'partitioning strategy' revealed here was ultimately used to direct upcoming sequential motor behavior. In terms of the small number of cells that encoded information about both targets, it is interesting to speculate whether they may provide an important 'bridge' between distinct motor plan elements or a higher conceptual representation of specific motor combinations not provided by the other subpopulations of neurons.

A concurrent BMI for planned sequential motor behavior

We exploited the simultaneous encoding and the neural partitioning mechanism observed in these experiments to develop a novel BMI *functionality* for performance of planned sequential motor behavior. This is a fundamental behavior in which all targets of a movement sequence are planned ahead of its initiation, and is largely distinct from behaviors involving the performance of independent single-targeted movements. The BMI functionality takes advantage of the concurrent encoding of a sequential motor plan in the premotor cortex, allowing it to determine all elements of the sequence simultaneously, upfront, and in advance of movement.

In addition, because the full motor plan is simultaneously decoded upfront and in advance of movement, the higher-level goal of the task can, in principle, also be analyzed before execution, and the motor plan reformulated accordingly. This could allow the prospective design of BMIs that can improve the performance of a sequential motor task, for example

perform the task more quickly, more flexibly, or more efficiently than originally conceived. Such a BMI may, for example, alter the order in which the elements of the motor sequence are executed depending on rapid or unpredictable changes in the environment (e.g., to avoid unanticipated obstacles), or correct the original sequence based on the performance metrics of the task (e.g., proactively change a sequence of letters based on spelling rules). As a simple but illustrative example of such a prospective capability (in the context of our experiments and using a relatively small number of recorded neurons), we demonstrate that we could accurately decode the full sequence of two targets in a very short time period after target presentation (Figs. 3, 4). Taken together, we demonstrate a concurrent BMI that allows the performance of a sequential motor behavior in line with how we naturally plan and execute it. Moreover, since information about all elements of the sequence is known ahead of execution, considering such concurrent decoding provides the future prospect of designing BMIs that can perform such tasks more effectively.

ONLINE METHODS

Behavioral task

Two adult rhesus monkeys (*Macaca Mulatta*) were trained to perform a working memory sequential delayed motor task. Monkeys were first sequentially presented with two distinct target locations on a screen, which were randomly selected in each trial, and then had to move a cursor to each in order by using a joystick (Dual-target task; Fig. 1a). After initial presentation of a blank screen, two targets were sequentially presented each of which could randomly take on one of four possible spatial locations: “up” (U), “down” (D), “right” (R) or “left” (L). To ensure that the two target locations were distinct, the motor sequence was chosen at random from a total of 12 possible sequences, i.e., all possible combinations of the two target locations excluding the ones with repeated locations. Targets were displayed for 500 ms each and were interleaved by a 300 ms interval during which a blank screen was shown. Following the end of second target presentation, there was an additional blank screen variable delay of 550– 850 ms (the working memory period) following which the first “go” cue signal appeared. After this, the monkeys were required to move a cursor from the center of the screen to the first remembered target. After reaching the target, they were required to return the joystick to the center and then wait for a second “go” cue to appear after an additional 500 ms delay interval. Once the second “go” cue appeared, they were allowed to move the cursor from the center of the screen to the second remembered target. The monkeys received a juice reward if they correctly moved to the two instructed targets.

Dual-target vs. single-target task

To examine the effect of adding information about a new target to working memory, it was necessary to disambiguate the process of holding information in working memory from that of adding information to it. To do this, primates performed randomly interleaved dual-target and single-target trials in a subset of sessions. On dual-target trials, described above, the primates were sequentially presented with two targets and then a blank screen delay. The time delay from the end of the first target presentation to the first “go” cue was therefore 1350–1650 ms. On single-target trials, in comparison, the primates were presented by only the first target, and had to keep this single target in working memory for the same total

1350–1650 ms time duration as in dual-target trials. However, here, they were not presented by a second target and were only shown a blank screen until the “go” cue.

Neurophysiologic recordings

All procedures were performed under IACUC-approved guidelines and were approved by the Massachusetts General Hospital institutional review board. Prior to recordings, multiple (up to six) planar silicone multi-electrode arrays (NeuroNexus Technologies Inc., MI) were surgically implanted in each monkey. Each of the implanted arrays contained four shanks horizontally spaced 400 μm apart. Every shank was 4 mm long and, in turn, contained 8 electrode contacts each vertically spaced 200 μm apart for a total of 32 contacts per electrode array. Hence the electrode contacts themselves spanned the bottom 1.6 mm of the shank. We advanced the electrodes approximately 2mm in depth. The electrode arrays were inserted into the cortex manually using microscope magnification. A craniotomy was placed over the premotor cortex under stereotactic guidance (David Kopf Instruments, CA). The multi-electrode arrays were separately implanted into the dorsal premotor (PMd) and the supplementary motor (SMA) areas (Supplementary Fig. 10). The electrode lead of each array was secured to the skull and attached to female connectors with the aid of titanium miniscrews and dental acrylic. Confirmation of electrode positions was done in both monkeys by direct visual inspection of the sulci and gyral pattern through the craniotomy. Additional post-mortem confirmation of electrode positions was made in one monkey (the second monkey is still performing experiments). Recordings began two weeks following surgical recovery. A Plexon multichannel acquisition processor was used to amplify and band-pass filter the neuronal signals (150 Hz – 8 kHz; 1 pole low-cut and 3 pole high-cut with 1000x gain; Plexon Inc., TX).

Shielded cabling carried the signals from the electrode array to a set of six 16-channel amplifiers. Signals were then digitized at 40 kHz and processed to extract action potentials in real time by the Plexon workstation. Classification of the action potential waveforms was performed using template matching and principle component analysis based on waveform parameters. Only single, well-isolated units with identifiable waveform shapes and adequate refractory periods (less than 1% of spikes within a 1 millisecond interval) were used for the online experiments and offline analysis. No multiunit activity was used.

Model construction

For the analysis of standard recording sessions, we model the activity of each neuron under any given sequence as an inhomogeneous Poisson process whose likelihood function is given by[49, 50]

$$p\left(N^{c}_{1:K}|S_i\right)=\prod_{k=1}^{K}\left(\lambda_c\left(k|S_i\right)\Delta\right)^{N^{c}_{k}}exp\left(-\lambda_c\left(k|S_i\right)\Delta\right)\quad i=1{:}12, \quad (1)$$

where Δ is the time increment taken to be small enough to contain at most one spike, N^{c}_{k} is the binary spike event of the c ’th neuron in the time interval [(*k* −1)Δ, *k*Δ], $\lambda_c\left(k|S_i\right)$ is its instantaneous firing rate in that interval, i S is the i ’th sequence, and K is the total number of bins in a duration KΔ . We take Δ = 5 ms as the bin width of the spikes. By building the

neuronal models under each sequence separately in the dual-target task we avoid making any a priori assumptions about whether the cells encode individual targets or combined sequences. For each sequence and neuron, we need to estimate the firing rate $\lambda_c(k|S_i)$ using the neuronal data observed. To do so, we use a state-space approach using the expectation-maximization (EM) algorithm[42, 51, 52] (see Supplementary Modeling). After fitting the models, we validated them using the χ^2 goodness-of-fit test on the data[42] and confirmed that they fitted the data well ($P > 0.7$ for all cells in all sessions).

Maximum-likelihood decoder

Once models are fitted, a maximum-likelihood decoder is used to decode the intended sequence based on the neuronal activity in any period of interest. A maximum-likelihood decoder is the optimal decoder in the sense of maximizing accuracy, i.e., the percentage of trials in which the combined sequence is decoded correctly, when the sequences are equally likely to be presented as is the case in our experiments. The decoder finds the likelihood of observing the population neuronal data under each sequence and selects the sequence with the highest likelihood as its prediction. Using the likelihood model in (1) and assuming that neurons are conditionally independent given the sequence, the population likelihood under any sequence is given by

$$p\left(N_{1:K}^{1:C}|S_i\right)=\prod_{c=1}^{C}\prod_{k=1}^{K}\left(\lambda_c\left(k|S_i\right)\Delta\right)^{N_k^c}exp\left(-\lambda_c\left(k|S_i\right)\Delta\right)\quad i=1{:}12,$$

where K is the total number of bins in any period of interest during the trial, C is the total number of neurons, and $\lambda_c(k|S_i)$ for $k = 1,\ldots,K$ and $c = 1,\ldots,C$ is the estimate of the firing rate. The predicted sequence, $\hat{S}$, is thus given by

$$\hat{S}=\underset{S_i}{arg\,max}\,p\left(N_{1:K}^{1:C}|S_i\right).$$

To find the sequence decoding accuracy of a single cell, the maximum-likelihood decoder uses only that cell's spiking activity to decode the sequence (Fig. 5 and Supplementary Fig. 5). The decoder also outputs the posterior probability of each sequence, which is the probability that it is the correct one after the neuronal observations, i.e.,

$$p\left(S_i|N_{1:K}^{1:C}\right)=\frac{p\left(N_{1:K}^{1:C}|S_i\right)}{\sum_i p\left(N_{1:K}^{1:C}|S_i\right)}\quad i=1,\cdots,12.$$

To dissociate the decoding accuracy of the first and second targets, denoted by T_1 and T_2, the decoder also outputs their predictions based on the neuronal activity. To do so, the decoder finds their posterior probabilities, i.e., $p\left(T_1=l_1|N_{1:K}^{1:C}\right)$ and $p\left(T_2=l_2|N_{1:K}^{1:C}\right)$, for all possible spatial locations, l_1 and l_2, by summing over the posterior probability of the sequences that have these spatial locations as their first or second targets. The decoder then

picks the spatial location with the highest first target (second target) posterior, or equivalently likelihood in our design, as its first (second) target prediction.

Comparison of the first target decoding accuracies in the single-target and dual-target tasks

To find the first target decoding accuracy of the single-target task we modeled the activity of each neuron under any given single target location as an inhomogeneous Poisson process, which was fitted using the EM procedure. We then performed the maximum-likelihood decoding analysis using leave-one-out cross-validation on the single-target trials. To make the comparison, for the dual-target task we constructed two models one for the first target and one for the second target, and then performed the decoding analysis for each target separately.

Determining the number of cells required to achieve the population accuracy

We found the number of cells required to achieve a given percentage of the population accuracy by first sorting them in each session based on their single neuron sequence accuracies and then performing the decoding analysis in that session for different number of cells in descending order.

BMI model training

In each BMI recording session, the monkeys first performed the dual-target task using a joystick (training session) during which models were constructed for the neuronal activity during an 800 ms time window prior to presentation of the “go” cue. This window length was chosen because in the standard dual-target sessions, it was sufficient to achieve better than 95% of the (maximum) sequence accuracy possible when using the entire window starting from second target presentation until the “go” cue (Supplementary Fig. 1b). We modeled the activity of each neuron in this window under any sequence as a homogeneous Poisson process (point process with constant rate), instead of an inhomogeneous one, to make the model construction faster for the BMI experiments. Hence using (1) the likelihood function for the spiking activity of neuron *c* under any of the sequences, S_i, was modeled as[49, 50]

$$p\left(N^c_{1:K}|S_i\right) = \prod_{k=1}^{K} \left(\lambda_c\left(S_i\right)\Delta\right)^{N^c_k} exp\left(-\lambda_c\left(S_i\right)\Delta\right),$$

where $\lambda_c(S_i)$ denotes the fitted firing rate of that neuron in the 800 ms window for sequence S_i and $K = 800/\Delta$ is the total number of bins in this period with bin width $\Delta = 5$ms. The firing rates were fitted using maximum likelihood parameter estimation. Here, the task involved either four (both monkeys) or eight (monkey P) sequences. The four-sequence task consisted of either “U-R”, “U-L”, “D-R”, “D-L” or “L-U”, “L-D”, “R-U”, “R-D”. The eight-sequence task consisted of the union of the sequences in the two four-sequence tasks.

The training sessions were followed by the real-time BMI sessions in which these trained Poisson models were used to predict the sequence using the maximum-likelihood decoder.

Concurrent online predictions and movement execution in the BMI

After the training sessions, the monkeys performed the same task as before. However, this time, cursor position was controlled by target predictions made by the maximum-likelihood decoder rather than the joystick. During the real-time BMI experiments, individual spike timings of all cells within the population were continuously recorded at a 40 kHz resolution by the Plexon multi-channel acquisition processor. Each recorded spike was then transmitted via an Ethernet port to a separate computer running a Matlab routine in real time. For each real-time trial, the Matlab routine then used the maximum-likelihood decoder to calculate the likelihood of the population spiking activity during the 800 ms time window prior to the "go" cue, i.e., $N_{1:K}^{1:C}$ under each sequence, S_i. This likelihood was calculated based on the trained Poisson models and assuming neurons were independent conditioned on the sequence. Hence the population likelihood for each sequence was found as

$$p\left(N_{1:K}^{1:C}|S_i\right) = \prod_{c=1}^{C}\prod_{k=1}^{K}\left(\lambda_c\left(S_i\right)\Delta\right)^{N_k^c} exp\left(-\lambda_c\left(S_i\right)\Delta\right)$$

The maximum likelihood decoder then outputted the sequence under which the population likelihood was maximized as the decoded sequence.

Based on the sequence decoded, a second Matlab routine running on the same computer then activated an analog output channel on the NI DAQ I/O interface to go from 0V to either +5V or −5V for 500 ms. The voltage line was, in turn, connected to a second NI DAQ I/O input channel located on a third computer running the behavioral program. Depending on the voltage received, the cursor displayed in the middle of the screen moved in a straight line to one of the four possible target locations (e.g., +5V in I/O channel 1 corresponded to a cursor location within the top target). This process then repeated for the second decoded target after another artificially introduced time delay. Here, the time delays in the two generated movements were selected to be similar to those that the monkeys normally experienced when performing the standard task using a joystick. However, the NI DAQ could in principle generate the two movements in as little as a few milliseconds apart.

Behavior versus prediction errors

Since the primates did not perform the dual-target working-memory task with 100% behavioral accuracy, some of the BMI errors were due to behavioral errors (i.e., the monkey not remembering the correct sequence during working memory) as opposed to decoder errors. Hence a more relevant accuracy number for the performance of the BMI could be the sequence accuracy obtained during the training session using leave-one-out cross-validation. This is because in the cross-validation analysis we calculate the accuracy by comparing the decoded sequence with the sequence the monkeys actually select after the "go" cue. For the BMI sessions, however, we compare the decoded sequence to the instructed sequence to find the accuracy. We hence tested whether after taking into account the primates' natural error rates, the accuracy during training sessions would be close to the BMI accuracy. Denoting the behavioral accuracy of the monkeys by P_b and the decoder accuracy found from the training session by P_t we can calculate what the accuracy of selecting the instructed

sequence would be after taking into account the behavioral errors. Denoting the resulting accuracy by P_f we have that

$$P_f = P_b P_t + (1 - P_b)(1 - P_t) \times \frac{1}{S - 1}.$$

In other words when the monkey and the decoder are both correct, the instructed sequence is selected. However, if the monkey is incorrect and the decoder is also incorrect in decoding the monkey's intended sequence, the probability of the decoder selecting the correct instructed sequence by random chance is $1 / (S - 1)$. We can find the mean and s.e.m. of P_f from those of P_b and P_t assuming P_b and P_t are independent[53] and then compare it with the BMI accuracy.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGEMENTS

R.C.U is funded by the NREF, E.N.B is funded by NIH DP1 OD003646, Z.M.W is funded by NIH 5R01-HD059852, PECASE and the Whitehall Foundation.

References

1. Chapin JK, Moxon KA, Markowitz RS, Nicolelis MAL. Real-time control of a robot arm using simultaneously recorded neurons in the motor cortex. Nat. Neurosci. 1999; 2:664–670. [PubMed: 10404201]
2. Wessberg J, et al. Real-time prediction of hand trajectory by ensembles of cortical neurons in primates. Nature. 2000; 408:361–365. [PubMed: 11099043]
3. Serruya MD, Hatsopoulos NG, Paninski L, Fellows MR, Donoghue JP. Instant neural control of a movement signal. Nature. 2002; 416:141–142. [PubMed: 11894084]
4. Hochberg LR, et al. Neuronal ensemble control of prosthetic devices by a human with tetraplegia. Nature. 2006; 442:164–171. [PubMed: 16838014]
5. Carmena JM, et al. Learning to control a brain-machine interface for reaching and grasping by primates. PLoS Biol. 2003; 1:193–208.
6. Taylor DM, Tillery SIH, Schwartz AB. Direct cortical control of 3D neuroprosthetic devices. Science. 2002; 296:1829–1832. [PubMed: 12052948]
7. Ganguly K, Carmena JM. Emergence of a stable cortical map for neuroprosthetic control. PLoS Biol. 2009; 7:1–13.
8. Wolpaw JR, McFarland DJ. Control of a two-dimensional movement signal by a noninvasive brain-computer interface in humans. Proc. Natl. Acad. Sci. USA. 2004; 101:17849–17854. [PubMed: 15585584]
9. Velliste M, Perel S, Spalding MC, Whitford AS, Schwartz AB. Cortical control of a prosthetic arm for self-feeding. Nature. 2008; 453:1098–1101. [PubMed: 18509337]
10. Moritz CT, Perlmutter SI, Fetz EE. Direct control of paralysed muscles by cortical neurons. Nature. 2008; 456:639–643. [PubMed: 18923392]
11. Mulliken GH, Musallam S, Andersen RA. Decoding trajectories from posterior parietal cortex ensembles. J. Neurosci. 2008; 28:12913–12926. [PubMed: 19036985]
12. Kim S-P, Simeral JD, Hochberg LR, Donoghue JP, Black MJ. Neural control of computer cursor velocity by decoding motor cortical spiking activity in humans with tetraplegia. J. Neural Eng. 2008; 5:455–476. [PubMed: 19015583]

13. Li Z, et al. Unscented Kalman filter for brain-machine interfaces. PLoS ONE. 2009; 4:1–18.
14. Chase SM, Schwartz AB, Kass RE. Bias, optimal linear estimation, and the differences between open-loop simulation and closed-loop performance of spiking-based brain-computer interface algorithms. Neural Networks. 2009; 22:1203–1213. [PubMed: 19502004]
15. Musallam S, Corneil BD, Greger B, Scherberger H, Andersen RA. Cognitive control signals for neural prosthetics. Science. 2004; 305:258–262. [PubMed: 15247483]
16. Santhanam G, Ryu SI, Yu BM, Afshar A, Shenoy KV. A high-performance brain-computer interface. Nature. 2006; 442:195–198. [PubMed: 16838020]
17. Shanechi MM, Wornell GW, Williams ZM, Brown EN. Feedback-controlled parallel point process filter for estimation of goal-directed movements from neural signals. IEEE Trans. Neural Syst. Rehabil. Eng. In press.
18. Shanechi, MM.; Williams, ZM.; Wornell, GW.; Brown, EN. A brain-machine interface combining target and trajectory information using optimal feedback control. Computational and Systems Neuroscience (COSYNE) Meeting; Salt Lake City, USA. 2011.
19. Kurata K. Premotor cortex of monkeys: Set- and movement-related activity reflecting amplitude and direction of wrist movements. J. Neurophysiol. 1993; 69:187–200. [PubMed: 8433130]
20. Messier J, Kalaska JF. Covariation of primate dorsal premotor cell activity with direction and amplitude during a memorized-delay reaching task. J. Neurophysiol. 2000; 84:152–165. [PubMed: 10899193]
21. Crammond DJ, Kalaska JF. Modulation of preparatory neuronal activity in dorsal premotor cortex due to stimulus-response compatibility. J. Neurophysiol. 1994; 71:1281–1284. [PubMed: 8201421]
22. Boussaoud D, Bremmer F. Gaze effects in the cerebral cortex: reference frames for space coding and action. Exp. Brain Res. 1999; 128:170–180. [PubMed: 10473755]
23. Crammond DJ, Kalaska JF. Differential relation of discharge in primary motor cortex and premotor cortex to movements versus actively maintained postures during a reaching task. Exp. Brain Res. 1996; 108:45–61. [PubMed: 8721154]
24. Crammond DJ, Kalaska JF. Prior information in motor and premotor cortex: Activity during the delay period and effect on pre-movement activity. J. Neurophysiol. 2000; 84:986–1005. [PubMed: 10938322]
25. Crutcher MD, Russo GS, Ye S, Backus DA. Target-, limb-, and context-dependent neural activity in the cingulate and supplementary motor areas of the monkey. Exp. Brain Res. 2004; 158:278–288. [PubMed: 15365665]
26. Hocherman S, Wise SP. Effects of hand movement path on motor cortical activity in awake, behaving rhesus monkeys. Exp. Brain Res. 1991; 83:285–302. [PubMed: 2022240]
27. Batista AP, Andersen RA. The parietal reach region codes the next planned movement in a sequential reach task. J. Neurophysiol. 2001; 85:539–544. [PubMed: 11160491]
28. Ninokura Y, Mushiake H, Tanji J. Representation of the temporal order of visual objects in the primate lateral prefrontal cortex. J. Neurophysiol. 2003; 89:2868–2873. [PubMed: 12740417]
29. Tanji J, Shima K. Role for supplementary motor area cells in planning several movements ahead. Nature. 1994; 371:413–416. [PubMed: 8090219]
30. Shima K, Isoda M, Mushiake H, Tanji J. Categorization of behavioural sequences in the prefrontal cortex. Nature. 2007; 445:315–318. [PubMed: 17183266]
31. Shima K, Tanji J. Neuronal activity in the supplementary and presupplementary motor areas for temporal organization of multiple movements. J. Neurophysiol. 2000; 84:2148–2160. [PubMed: 11024102]
32. Baldauf D, Cui H, Andersen RA. The posterior parietal cortex encodes in parallel both goals for double-reach sequences. J. Neurosci. 2008; 28:10081–10089. [PubMed: 18829966]
33. Averbeck BB, Sohn J-W, Lee D. Activity in prefrontal cortex during dynamic selection of action sequences. Nat. Neurosci. 2006; 9:276–282. [PubMed: 16429134]
34. Mushiake H, Saito N, Sakamoto K, Itoyama Y, Tanji J. Activity in the lateral prefrontal cortex reflects multiple steps of future events in action plans. Neuron. 2006; 50:631–641. [PubMed: 16701212]

35. Ohbayashi M, Ohki K, Miyashita Y. Conversion of working memory to motor sequence in the monkey premotor cortex. Science. 2003; 301:233–236. [PubMed: 12855814]
36. Kettner RE, Marcario JK, Port NL. Control of remembered reaching sequences in monkey. II. Storage and preparation before movement in motor and premotor cortex. Exp. Brain Res. 1996; 112:347–358. [PubMed: 9007537]
37. Lu X, Ashe J. Anticipatory activity in primary motor cortex codes memorized movement sequences. Neuron. 2005; 45:967–973. [PubMed: 15797556]
38. Nakajima T, Hosaka R, Mushiake H, Tanji J. Covert representation of second-next movement in the pre-supplementary motor area of monkeys. J. Neurophysiol. 2009; 101:1883–1889. [PubMed: 19164110]
39. Averbeck BB, Chafee MV, Crowe DA, Georgopoulos AP. Parallel processing of serial movements in prefrontal cortex. Proc. Natl. Acad. Sci. USA. 2002; 99:13172–13177. [PubMed: 12242330]
40. Saito N, Mushiake H, Sakamoto K, Itoyama Y, Tanji J. Representation of immediate and final behavioral goals in the monkey prefrontal cortex during an instructed delay period. Cereb. Cortex. 2005; 15:1535–1546. [PubMed: 15703260]
41. Mushiake H, Inase M, Tanji J. Selective coding of motor sequence in the supplementary motor area of the monkey cerebral cortex. Exp. Brain Res. 1990; 82:208–210. [PubMed: 2257906]
42. Smith AC, et al. State-space algorithms for estimating spike rate functions. Computational Intelligence and Neuroscience. 2010; 2010
43. Tanji J. Sequential organization of multiple movements: Involvement of cortical motor areas. Annu. Rev. Neurosci. 2001; 24:631–651. [PubMed: 11520914]
44. Cisek P, Kalaska JF. Neural correlates of reaching decisions in dorsal premotor cortex: Specification of multiple direction choices and final selection of action. Neuron. 2005; 45:801–814. [PubMed: 15748854]
45. Hoshi E, Tanji J. Integration of target and body-part information in the premotor cortex when planning action. Nature. 2000; 408:466–470. [PubMed: 11100727]
46. Hoshi E, Tanji J. Contrasting neuronal activity in the dorsal and ventral premotor areas during preparation to reach. J. Neurophysiol. 2002; 87:1123–1128. [PubMed: 11826076]
47. Hoshi E, Tanji J. Differential involvement of neurons in the dorsal and ventral premotor cortex during processing of visual signals for action planning. J. Neurophysiol. 2006; 95:3596–3616. [PubMed: 16495361]
48. Shima K, Tanji J. Both supplementary and presupplementary motor areas are crucial for the temporal organization of multiple movements. J. Neurophysiol. 1998; 80:3247–3260. [PubMed: 9862919]
49. Brown, EN.; Barbieri, R.; Eden, UT.; Frank, LM. Likelihood methods for neural data analysis. In: Feng, J., editor. Computational Neuroscience: A Comprehensive Approach. CRC Press; Boca Raton: 2003. p. 253-286.
50. Truccolo W, Eden UT, Fellows MR, Donoghue JP, Brown EN. A point process framework for relating neural spiking activity to spiking history, neural ensemble, and extrinsic covariate effects. J. Neurophysiol. 2005; 93:1074–1089. [PubMed: 15356183]
51. Smith AC, Brown EN. Estimating a state-space model from point process observations. Neural Comput. 2003; 15:965–991. [PubMed: 12803953]
52. Dempster AP, Laird NM, Rubin DB. Maximum likelihood from incomplete data via the EM algorithm. J. Roy. Statist. Soc. B. 1977; 39:1–38.
53. Goodman LA. On the exact variance of products. J. Amer. Statistical Assoc. 1960; 55:708–713.

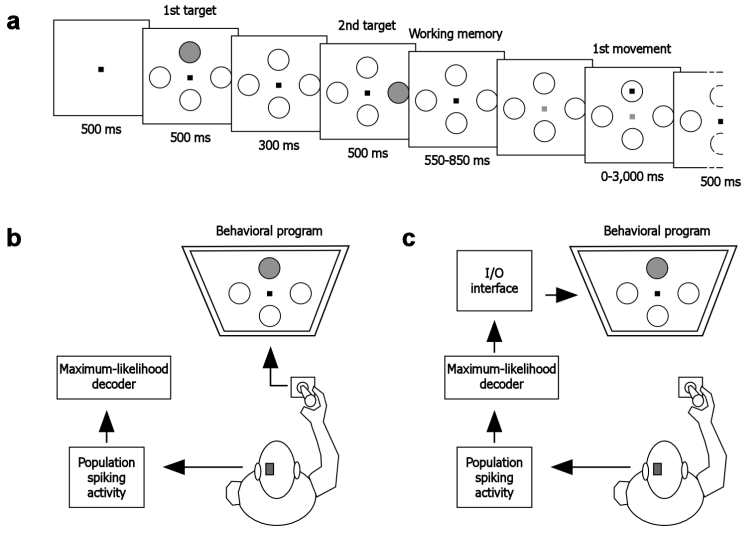

Figure 1.

Task design and experimental setup. (**a**) Schematic illustration of a standard dual-target task over a single trial. Task periods and their timings are displayed over a single trial from left to right. The right end of the panel in which the second movement is made is truncated to conserve space. Decoding analyses are performed during the 500 ms blank-screen interval following presentation of the second target. (**b**) Experimental setup for the standard training sessions. (**c**) Experimental setup for the BMI sessions.

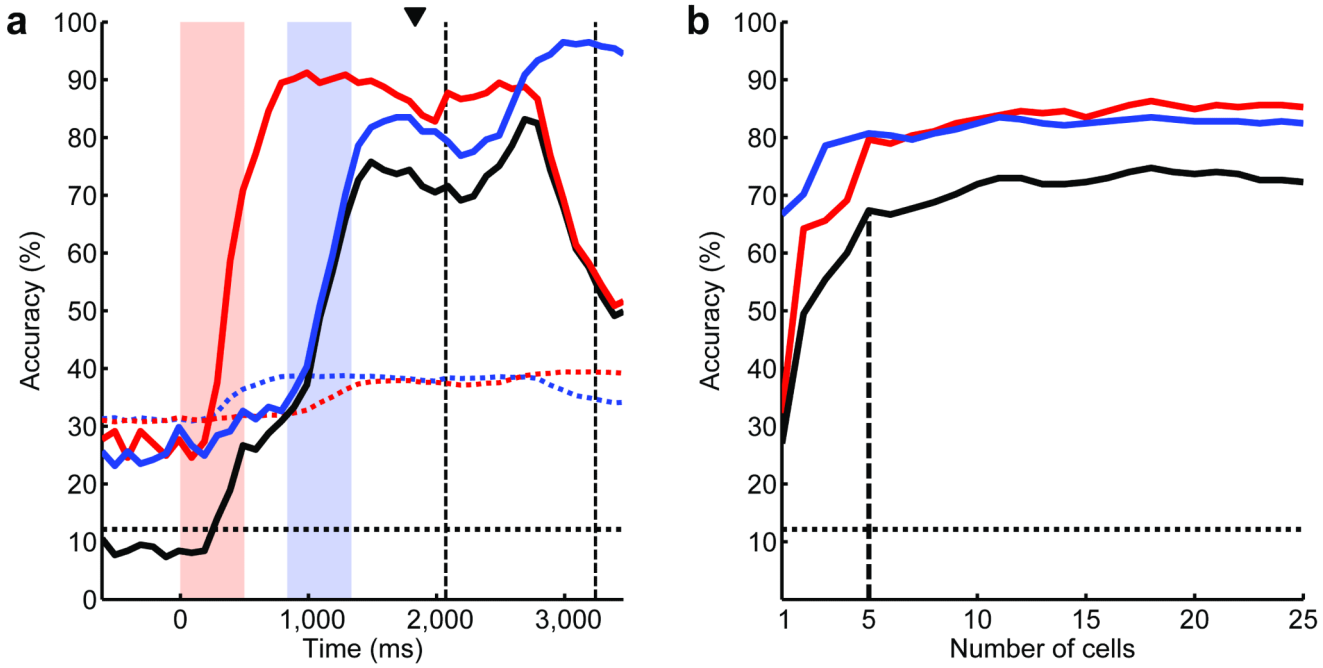

Figure 2.

Population decoding accuracy for a selected session. (**a**) Population decoding accuracy over time for the first target (red curve), second target (blue curve), and the full sequence (black curve). Each point on the curves indicates the decoding accuracy for the population over the *preceding* 500 ms window. Time at zero is aligned to the start of first target presentation. The red and blue vertical bars indicate the times during which the first and second targets were presented, respectively. The first and second dashed black lines indicate the mean times at which the first and second "go" cues were given, respectively. The arrow indicates the time point of decoding for the preceding working memory period (i.e., 0–500 ms from the end of the second target presentation). The dotted lines indicate the 99% chance upper confidence bounds for the first target, second target, and sequence (out of 12 possibilities), with the same respective color scheme used above (see also Supplementary Modeling). (**b**) Number of cells sufficient to reach decoding accuracy asymptote during the working memory period for the same session. The first target (red curve), second target (blue curve), and sequence (black curve) accuracies are displayed as a function of the cumulative number of cells, in descending order of single-cell sequence accuracy. The number of cells needed to reach over 90% of the population accuracy is indicated by the vertical dashed line.

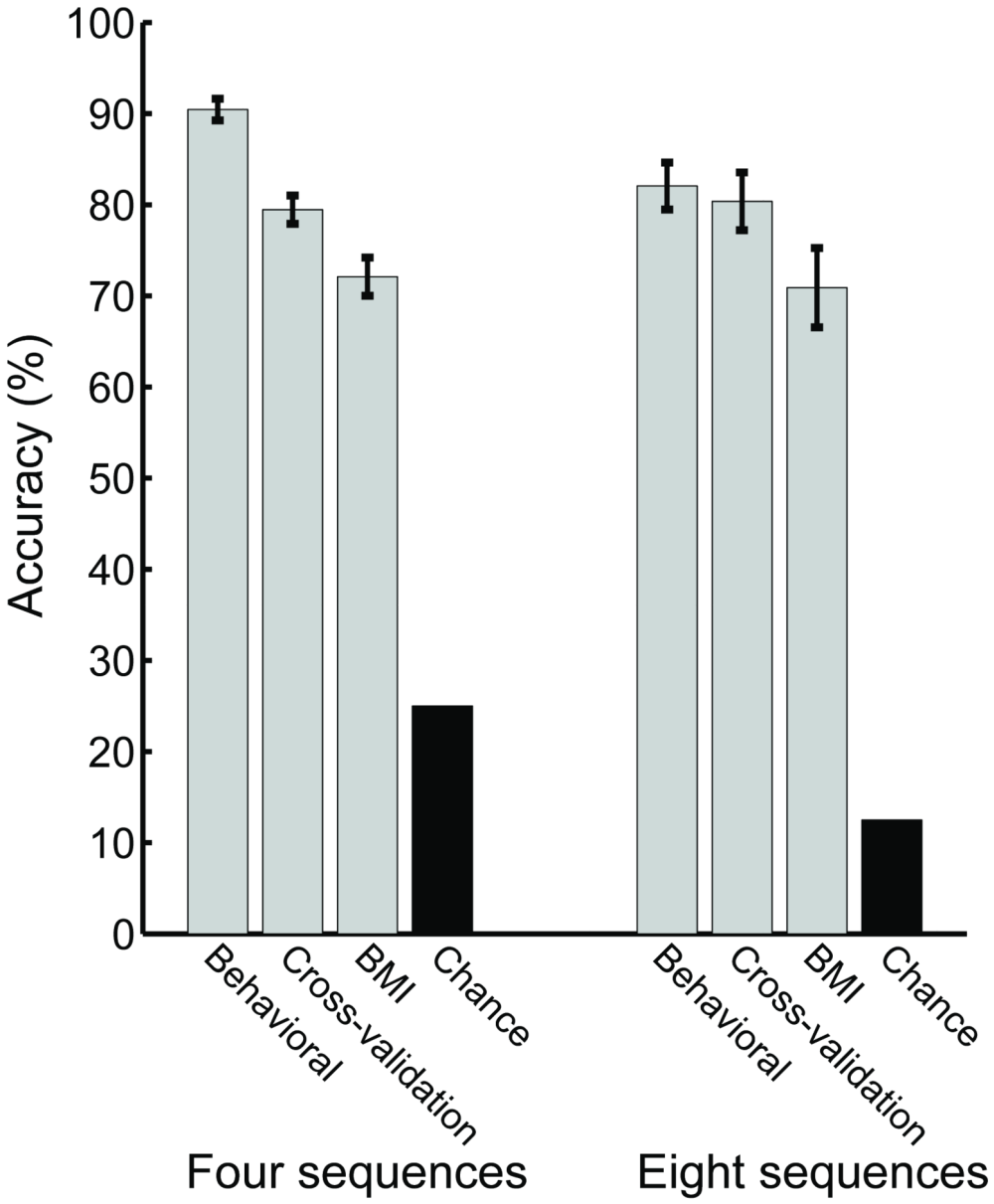

Figure 3.

Decoding accuracies on BMI trials. The gray bars indicate the monkeys' average behavioral accuracy, maximum-likelihood cross-validation accuracy on the training data, and real-time BMI accuracy, with their corresponding s.e.m.. The black bars indicate chance level accuracies. Performances using four sequences are displayed on the *left*, and using eight sequences are displayed on the *right*.

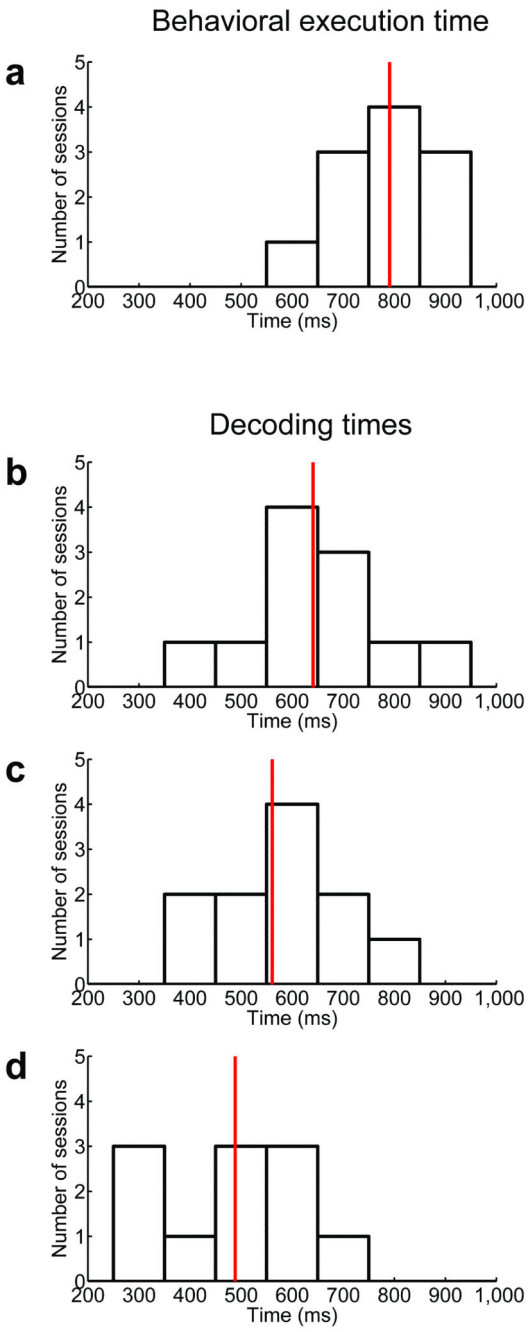

Figure 4.

Decoding and behavioral performance times. (a) Histogram of the total times it took the monkeys to both *behaviorally* react to the two "go" cues and moreover reach the two targets (excluding any task delays and the time required to move between targets). (b , c and d) Histograms of the times required for the *decoding* accuracy to reach 90% asymptotic accuracy, from the time of second target presentation, for 12, 8, and 4 sequences, respectively. The red line indicates the mean times for each histogram.

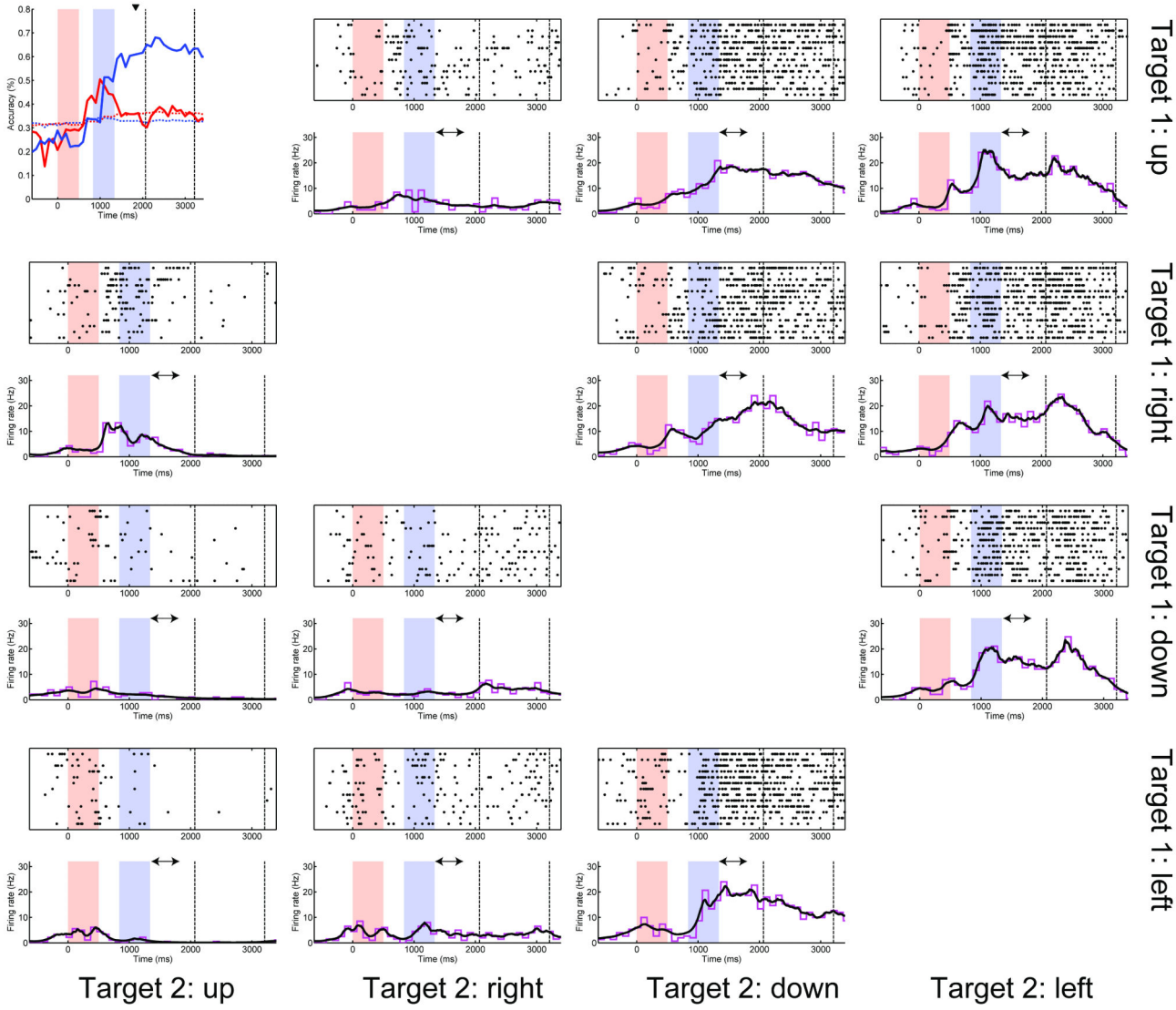

Figure 5.

Example of a second (added) target selective neuron. The subfigure at the upper left corner shows the first and second target accuracies of the cell as a function of time into the trial. The vertical bars/lines and their timings follow the same convention as Fig. 2. In all other subfigures, each *top* panel corresponds to a different sequence of movements with each row illustrating the spiking activity during a single trial and the black dots indicating the spike times. Each *bottom* panel indicates the corresponding mean firing rate estimates using the expectation-maximization procedure (black curve) and the corresponding peristimulus time histogram (PSTH) (magenta curve). The arrow indicates the working memory period. The subfigures in the same row correspond to sequences with the same first target location. The subfigures in the same column correspond to sequences with the same second target location. Note that repeated targets locations were not used in the sequences and hence there are 3 subfigures per row/column (see also Supplementary Figs. 3, 4).

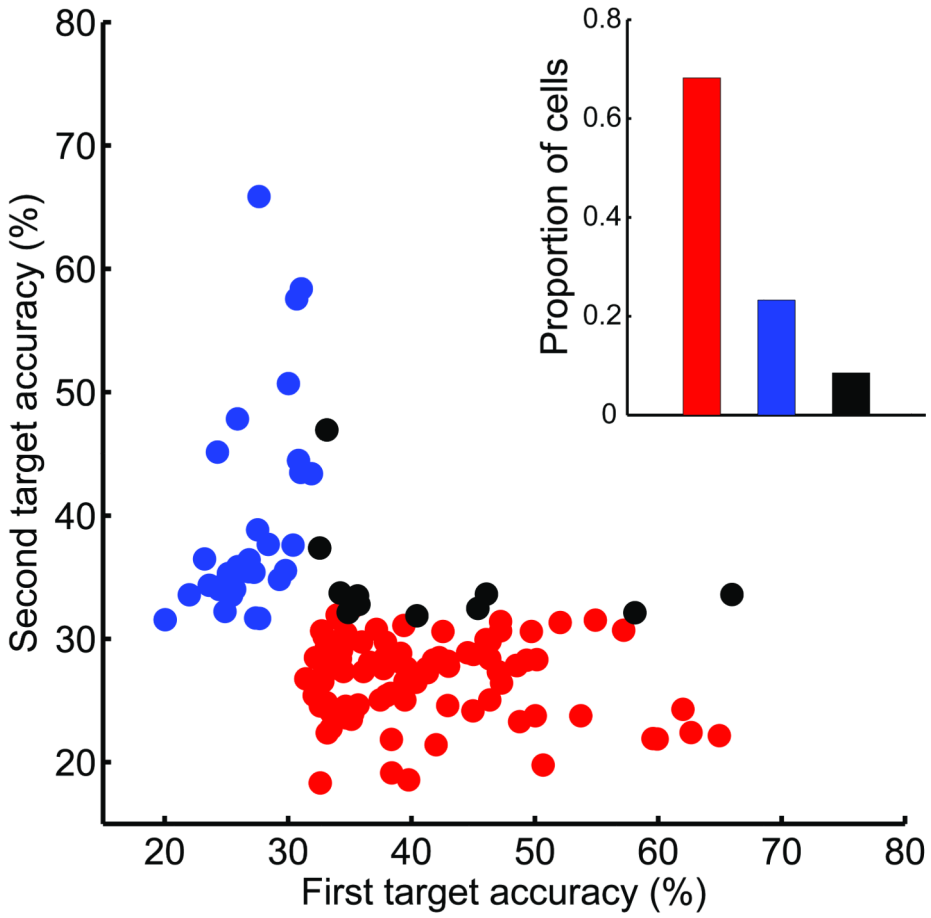

Figure 6.

Distribution of first and second target information across the population. Scatter plot of the first and second target accuracies for the 129 cells that encoded significant information about at least one target during the working memory period (across 12 sequences). Statistical significance of the target accuracies was tested here at the 0.01 level (see also Supplementary Fig. 6). Red points indicate cells that encoded significant information about only the first target, blue points indicate those that encoded significant information about only the second target, and black points indicate those that encoded significant information about both targets. The *inset* indicates the proportion of cells that encoded significant information about only the first, only the second or both targets during the working memory period with the same coloring scheme, above.

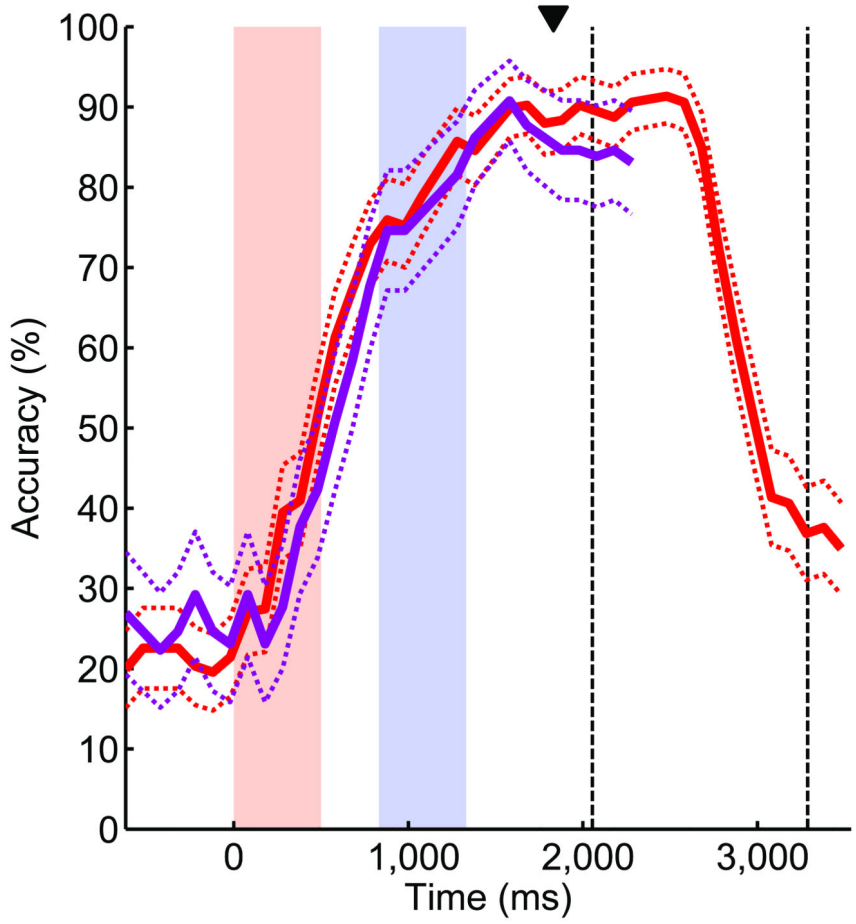

Figure 7.

The effect of adding information to working memory. In an interleaved session, population decoding accuracy for the first target on dual-target trials is shown in red whereas population decoding accuracy for the first target on single-target trials is shown in magenta. Each point on the curves indicates the decoding accuracy over the *preceding* 500 ms window. Dotted lines indicate the 95% confidence bounds for accuracy of each curve (rather than chance level). The vertical red bar indicates the time during which the first target was presented. The vertical blue bar indicates the time during which the second target was or was not shown depending on the trial type. The arrow indicates the time point corresponding to the decoding accuracy of the preceding working memory period (same convention as in Fig. 2).