"A High-Performance Brain Computer Interface" Supplementary Information

G. Santhanam, S.I. Ryu, B.M. Yu, A. Afshar, K.V. Shenoy

Contact: shenoy@stanford.edu

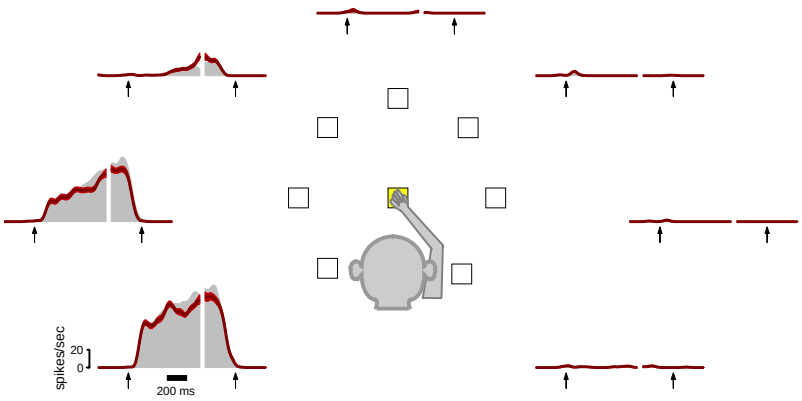
1 Video Legends

These videos show the behavioral conditions from our different experiments. The experiment room itself was visibly dark, but the video camera could image the scene with infrared light. Bright text and numbers were added as annotations and were not seen by the monkey. Each successful trial was followed by a juice reward, indicated by a short tone followed by a click corresponding to the action of a juice dispenser. These auditory signals were heard by the monkey during the experiment. For the duration of the experiment, a piece of reflective tape was lightly wrapped around one phalanx of the monkey's left hand to track the monkey's arm movements. This tape is especially bright in the infrared videos but was not visible to the monkey during the experiments.

Video 1 SuppVideo1RealReach.mpg Movie showing the instructed-delay task. The monkey began each trial by touching the center square and fixating the cross. The reach target appeared in the periphery and the monkey was required to wait until the 'go' cue (an extinguishing of the center cues as well as a slight enlargement of the reach target) before making the reach.

Video 2 SuppVideo2BCIModerate.mpg Movie showing moderate-paced prosthetic cursor trials during BCI experiments. Video annotations are used to denote real reaches and prosthetic cursor trials. Numbers indicate a particular trial's position in a contiguous sequence of cursor trials. For each cursor trial, we decoded the target location, and a prosthetic cursor (white circle) was rendered on screen. Real reach trials were interspersed to ensure the monkey remained engaged in the task.

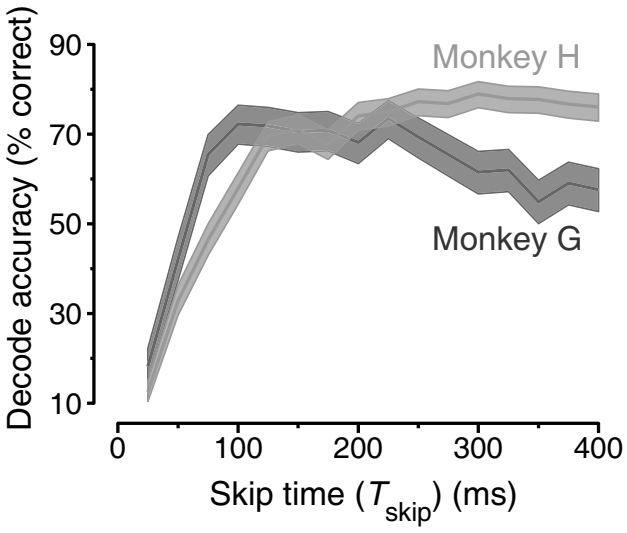
Video 3 SuppVideo3BCIFast.mpg Movie showing fast-paced prosthetic cursor trials during BCI experiments. We showed three separate sequences, each of five cursor trials followed by a real reach. These three sequences were not performed in succession. Rather, the three sequences of five cursor trials were spliced from different times during the same experiment to better demonstrate the speed of our system. The prosthetic cursor is not visible in this video because it was flashed briefly so as to not slow the overall presentation of trials.



Supplementary Figure S1: Neural response of unit `H20050329.25.1` in the conventional instructed-delay task (solid gray; height denotes mean response) and the blink-memory task (red; solid line denotes mean and shading represents ±1 SEM). In the blink-memory task, the target was shown for 200 ms at the start of the delay period and remained extinguished until the 'go' cue. Both tasks were interleaved in a pseudorandomized fashion during each experiment. Spiking data for each trial was filtered using a Gaussian function and data were averaged across trials. The first arrow designates the target presentation time and the second arrow marks the start of the movement. The delay and movement periods are separated by a slight gap to allow for differing delay periods across trials.

2 Motor-related activity in the absence of the visual stimulus

We performed a control experiment to confirm that sustained neural activity during the delay period was related to motor intention rather than to a primary sensory response. This experiment used a "blink-memory" task design where the reach target was blinked briefly and neural activity was measured during a memory period when there was no visual stimulus. At the time of the 'go' cue, the target was shown again and the monkey was requested to reach to this location. Supplementary Fig. S1 shows the neural activity from one neuron for each of 7 target locations. The neural response is compared between the conventional instructed-delay (solid gray) and blink-memory (red) tasks. The activity does not rely on the continued presence of a visual target. The neuron in Supp. Fig. S1 was typical. There were also neurons for which activity for the blink-memory task was elevated over that of the conventional instructed-delay task. In the cases where there was a difference in the activity, the difference was often only present for one or two target locations out of the seven possible locations. This result is consistent with prior publications (Crammond & Kalaska, 2000; Cisek & Kalaska, 2002).



Supplementary Figure S2: PMd latency analysis with the single-target instructed-delay task (one reach target was shown out of a possible of 8 locations and the remaining 7 locations are invisible) as a function of T_{skip}. Performance was calculated by training a Poisson model on all trials in a dataset and computing the leave-one-out cross-validated performance on the same data. The shaded area denotes the 95% confidence interval (Bernoulli process) around the mean performance (embedded line). Dark curves correspond to monkey G (dataset `G20040603`) and light curves to monkey H (dataset `H20041117`). Performance was calculated for a constant T_{int} of 50 ms with varying T_{skip}.

3 Selection of T_{skip}

The choice of T_{skip} affects the performance of our BCI. Because our target decoding models are constructed under the assumption that there is statistical stationarity across trials, it is important for T_{skip} to include the following time intervals: (1) the time for visual information about the target to arrive in PMd (50–70 ms), (2) the time for the subject to select among targets if more than one are present, and (3) the time for neural activity reflecting the desired target to be generated.

Before visual information is relayed to PMd, the measured neural activity in PMd is not target-related — random neural variability can inject noise into our decoding model. We computed single-trial accuracy as a function of T_{skip}, fixing T_{int} to 50 ms. Supplementary Fig. S2 demonstrates that the neural activity in PMd cannot be meaningfully decoded to predict the reach target until $\sim$75 ms after the target is displayed. This estimate includes a $\sim$16–33 ms delay between when the software sends a request to show the stimulus and when it is actually displayed by the CRT projector. Supplementary Fig. S2 also reveals that there is target related information in PMd as early as 50–70 ms after the target is first cued. It would not be possible to decode the target with

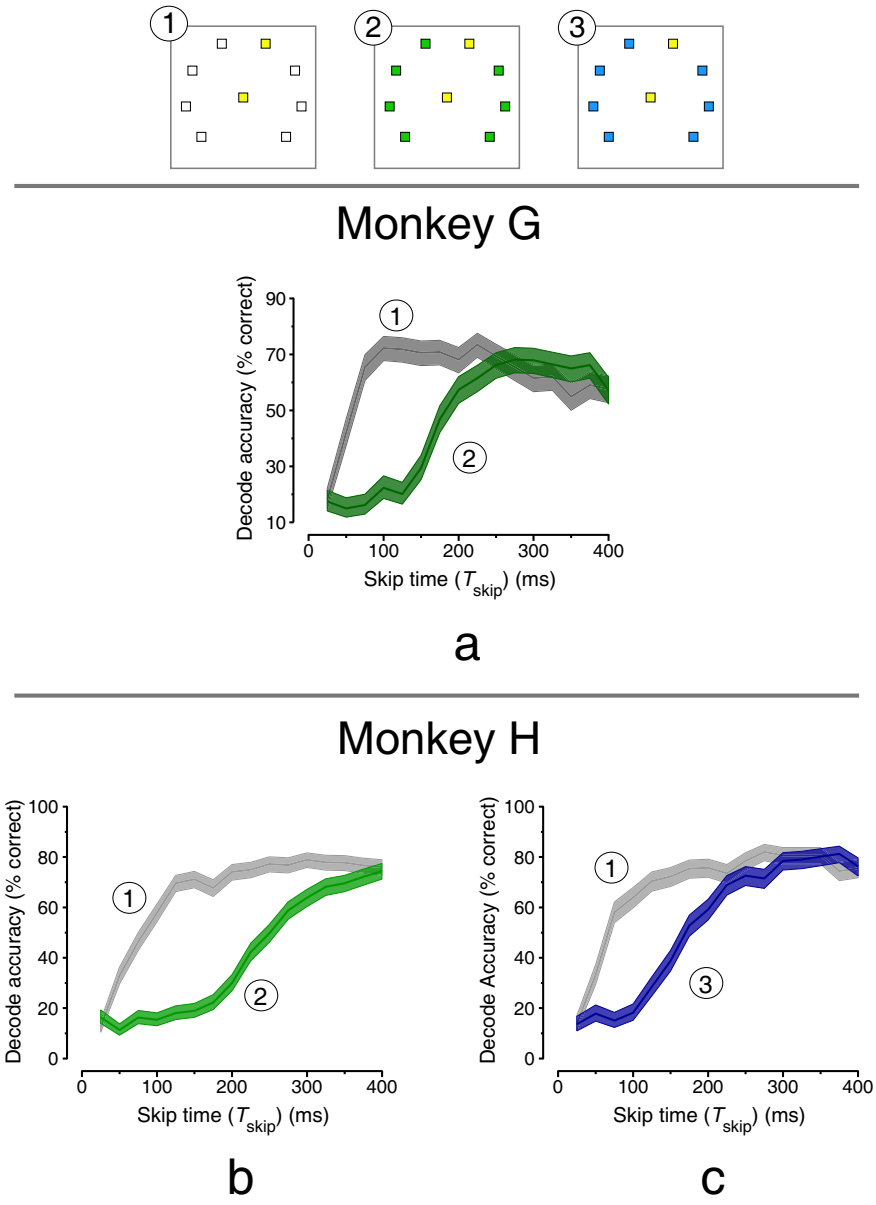
above chance probability otherwise. This rough estimate of latency agrees with neural response plots from other previous studies in PMd (Crammond & Kalaska, 2000; Kalaska & Crammond, 1995, etc.), where some neurons show a change in activity very soon after stimulus onset. *This exact latency has further implications for BCI experiments where reach targets are presented in rapid succession. Figure 1b of the main text shows that neurons were spiking according to the target location of a previous trial for many 10s of milliseconds after the start of a new trial (see just after ellipse #2).*

To estimate the time for the brain to select among multiple reach targets, we performed a separate control experiment with both monkeys. We presented the monkey with a multi-target task where all of the eight possible reach locations were shown on every trial, but only one was colored yellow while the rest were colored green. The monkey was trained to reach for the yellow target following the delay period. Supplementary Fig. S3 compares the performance for the conventional single-target instructed-delay and multi-target tasks, as a function of T_{skip}. For the multi-target task, we require a longer T_{skip} before there is a decodable reach plan. For monkey H, we used both a yellow-green and yellow-blue color scheme.[1] Comparing the two color schemes, there is a much larger (+150 ms) latency for the yellow-green scheme. This large difference between the two *color* schemes demonstrates that the difficulty of the task can greatly influence the speed at which plans are formed.

A question that frequently arises in visually cued studies such as ours is whether the neural activity measured during the delay period is related to a reach plan, the visually cued stimulus, or a combination of both. One such discussion of this issue can be found in Crammond & Kalaska (1995). For example, recording from primary visual cortex could provide excellent prospects for decoding the reach target in our single-target task, but a BCI operating on this neural activity would not represent the motor intentions of the subject. Supplementary Fig. S1 partly addresses this concern, and our multi-target task can also serve as a control experiment in this regard. Placing T_{skip} at the time where the performance curves in Supp. Figs. S3a and S3c converge would provide assurance that such a BCI is decoding motor intention.[2]

[1] All colors were measured to be roughly isoluminant via a photometer.

[2] Analyses similar to those of Supp. Fig. S3 demonstrated no drop in performance for the blink-memory task after the target was extinguished during the delay period.



Supplementary Figure S3: Direct performance comparison between the single-target and multi-target tasks as a function of T_{skip}. Both tasks were interleaved in a pseudorandomized fashion during each experiment. Analysis is similar to that presented in Supp. Fig. S2. Performance is plotted with T_{int} fixed at 50 ms and T_{int} varied. **a**. Performance with yellow-green color scheme converges at T_{skip}≈250 ms (dataset `G20040603`). **b**. Performance with yellow-green color scheme converges at T_{skip}≈400 ms (dataset `H20041117`). **c**. Performance with yellow-blue color scheme converges at T_{skip}≈250 ms (dataset `H20041201`).

The multi-target task is also an inherently more difficult task than the single-target task. The different time courses of these two tasks cannot be entirely due to the difference in visual stimuli, especially since merely changing the color of the *non-reach targets* caused a considerable shift in the time course for monkey H (cf. Supp. Figs. S3b and S3c). We therefore chose to be neither overly conservative by waiting until the time at which the performance curves fully converged (250 ms), nor overly liberal by selecting T_{skip} to be coincident with the early plateau in decoder accuracy for the single-target task (75–100 ms). We chose a T_{skip} of 150 ms.

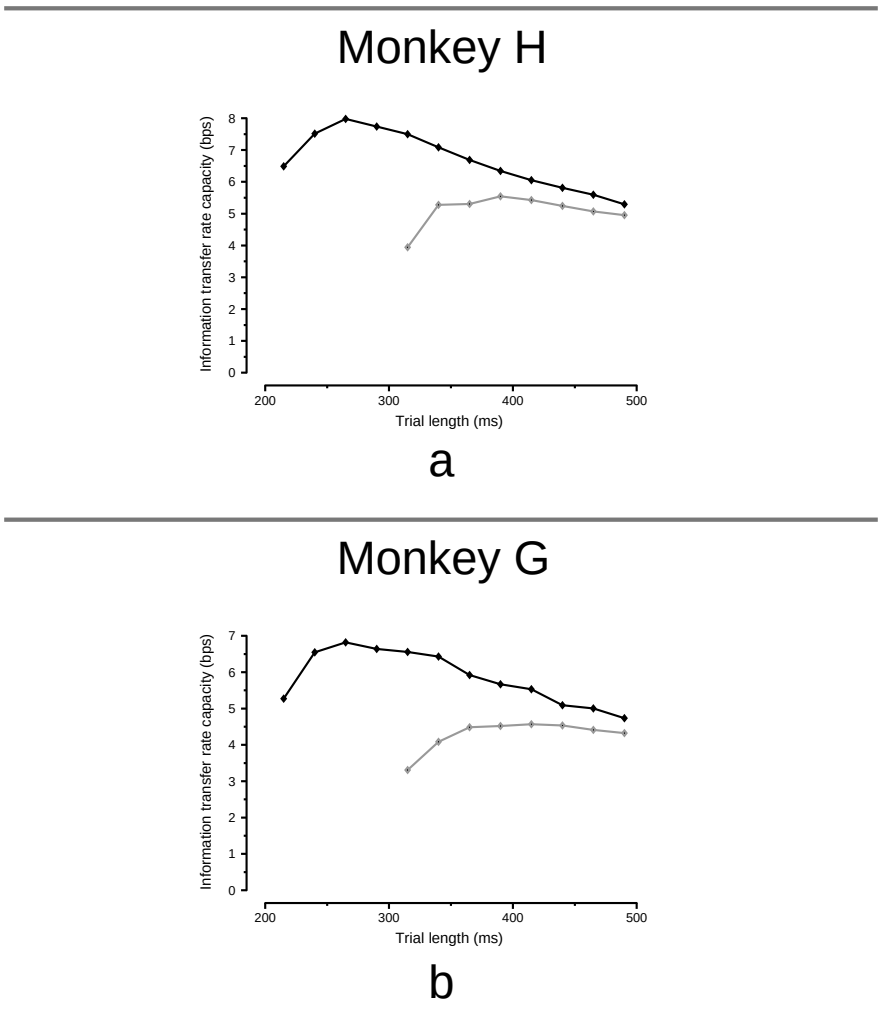
We used the single-target task, as opposed to the more complicated multi-target task, for our BCI experiments because we felt that it provided the simplest analogy to a human prosthetic system. While real patients may have to choose from several objects in their workspace, these objects will not ordinarily be presented immediately prior to a decision to execute a prosthetic reach. Furthermore, BCIs will typically rely on internally-generated target plans as opposed to externally presented stimuli. BCIs tested under with internally-generated plans may well achieve even greater performance than what we have demonstrated in this report. Internally generated plans can be formed without the added latencies of the visual system. Experiments are underway to test this hypothesis.

4 ITRC for multi-target task

Since the multi-target task is a more difficult task, overall performance of a BCI using such a paradigm may not be as high as that of the single-target-based system. Supplementary Fig. S4 compares the ITRC between the single-target task and the multi-target task in control experiments. In summary, there was a ∼30% penalty for a system using the more difficult multi-target task.[3] Similar to the discussion for T_{skip}, the difference in ITRC performance could be attributed to differences in visual stimulus presentation, cognitive difficulty, or a combination of both.

Presumably the performance degradation would be similar when running BCI experiments using a multi-target task. Adjusting our BCI performance of 6.5 bits/s by 30% yields 4.6 bits/s. This adjusted ITRC is still well in excess of previously reported results. Importantly almost all past studies are based on single, visual stimulus, including BCIs employing continuous trajectory control. This was one reason for why we chose to employ a single-target paradigm for BCI experiments and

[3]The maximum ITRC in this analysis differs from that found in Supp. Fig. 2 of the main text since different datasets were used for each analysis along with different model training methods.



Supplementary Figure S4: ITRC comparisons between single-target (black) and multi-target (gray) tasks. An 8-target layout was used in the experiment. Both tasks were interleaved in a pseudorandomized fashion during each experiment. Trial length was taken to be $T_{\mathsf{skip}}+T_{\mathsf{int}}+T_{\mathsf{dec+rend}}$with $T_{\mathsf{dec+rend}}$ set to 40 ms. T_{skip} was fixed at 150 ms for the single-target task and 250 ms for the multi-target task. T_{int} was varied and performance extrapolated. **a**. Data from monkey H (`H20041201`) with a Poisson decoding model; maximum ITRC was 8.0 bits/s for the single-target instructed-delay task and 5.5 bits/s for the multi-target task **b**. Data from monkey G (`G20040603`) with a Gaussian decoding model; maximum ITRC was 6.8 bit/s and 4.6 bit/s, respectively.

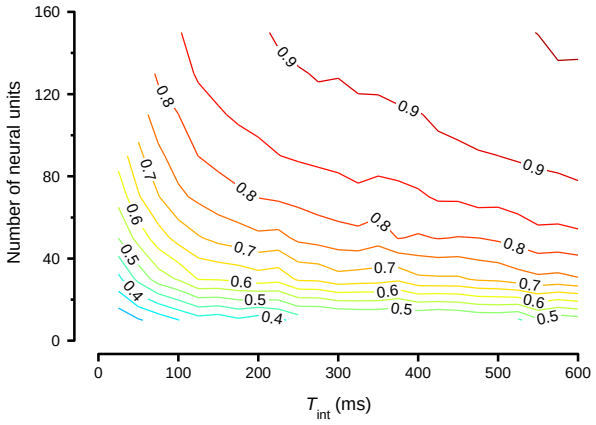
report those results in the main text.

5 ITRC comparisons between control and BCI experiments

As with monkey H, all BCI experiments with monkey G yielded ITRC values consistently less than those predicted by extrapolations in our control experiments. For example, the best ITRC obtained for this monkey was 5.3 bits/s at T_{int}=200 ms in our BCI experiments. For this same value of T_{int}, the extrapolations predict an ITRC of 6.4 bits/s, again supporting the idea that extrapolations from slower-paced trials will overestimate performance when targets are presented rapidly.

Performance in BCI experiments can fall short of extrapolated performance for a variety of reasons. One possible source of such declines includes situations where data used to train the decoding models is dissimilar to data used for prediction. To optimize the similarity between these two conditions, for monkey H we presented rapid sequences of reach targets during the training portion of our experiment, only commanding the monkey to reach for the last target in the sequence. Since the subject presumably planned reaches to every target, statistical models were trained from these high-speed trials that mimic the speed of trials during the prediction portion of the experiment. Overall performance was improved in comparison to decoding using models trained on slower-paced trials.

6 Effects of losing neural units on single-trial accuracy



Supplementary Figure S5: Single-trial accuracy as a function of numbers of units and T_{int}. All data is from experiment `H20041118` which involved an 8-target configuration. T_{skip} was fixed at 150 ms. Similar results were obtained for dataset `G20040508` from monkey G.

At a population subset of ∼80 neural units, increasing the subset size improves decode performance. Increasing the integration time also improves decode performance. For very few numbers of neural units, the performance eventually saturates regardless of the size of T_{int}. This effect reflects the inherent noise present from sampling a small subset of neurons as well as the potential mismatch of our (or any) spiking model to the actual statistics of the neural system.

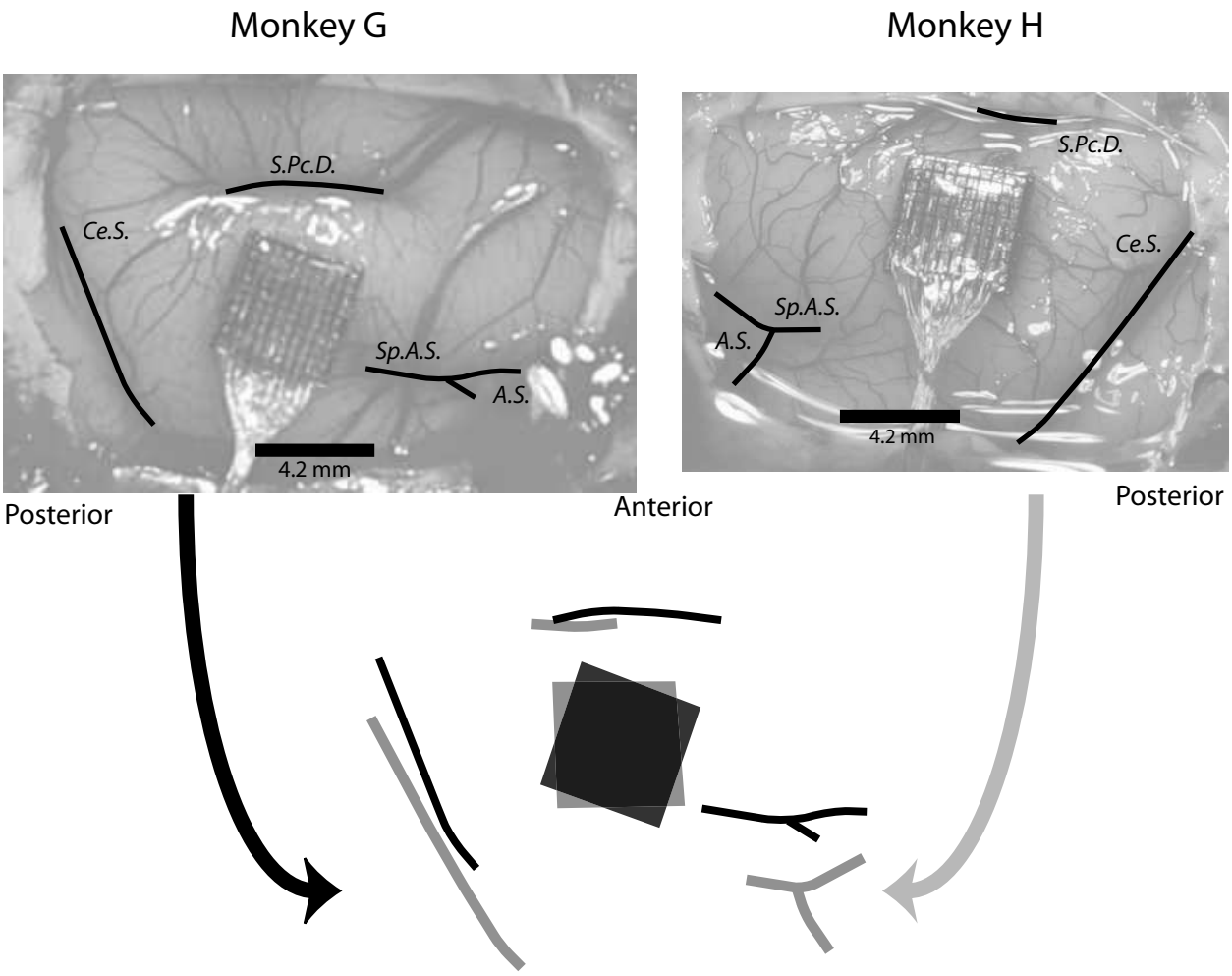
7 Supplementary Methods

Given the interleaved presentation of real reach and BCI trials, the subject was not aware at the start of a trial whether or not he will make a reach to the presented target. Neural activity in premotor cortex represents the potential reach target, but a conceptual 'gate' that connects this neural activity to muscle activity remains closed since movements were explicitly not permitted during the delay-period. For real reach trials, the subject was given and later recognized the 'go' cue, thereby opening the 'gate,' allowing the primed activity in premotor cortex (and other areas in the CNS such as the primary motor cortex, thalamus, and spinal cord) directed the arm movement. Other prosthetic studies have previously shown similar gating in a different situation. Experiments have demonstrated that prosthetic devices trained on with M1 and PMd neural activity during a subject's real arm movements can function even when the subject's arm is loosely restrained (Serruya et al., 2002; Taylor et al., 2002; Carmena et al., 2003), revealing that recorded neurons largely create similar activity regardless of whether or not the 'gate' to muscle activity is opened or closed. Furthermore, an fMRI study has shown that there is little change in the motor cortical activity of tetraplegics, even when the 'gate' has been left closed for several years due to paralysis (Shoham et al., 2001).

Neural Recordings

Supplementary Fig. S6 shows the anatomical placement of the electrode array in both monkeys.

In BCI experiments, a selection process determined which units were to be used during target prediction. For monkey G, we used a first-generation automated spike sorting infrastructure (Santhanam et al., 2004). Neural data was collected from all electrodes (immediately prior to conducting the actual experiment). This data was processed by several computers to determine the number of units on each electrode as well as the mean waveform for each unit, clustering the data in a modified principal components space (Sahani, 1999). The waveforms were then used to calculate



Supplementary Figure S6: Placement of electrode arrays in PMd of monkeys G and H. For both monkeys, the arrays were placed in a location that spans dorsal premotor and primary motor cortices. The neural signals tended to be responsive during both the delay period and the movement phase of trials. Intraoperative photographs of the array implanted in cerebral cortex are shown with sulci indicated. Overlapping diagram shows the relative array placement between monkeys. Monkey H's sulcal pattern is reflected vertically and rotated to bring the sulci into alignment with those of monkey G. Ce.S.: central sulcus; S.Pc.D.: superior precentral dimple; Sp.A.S.: spur of the arcuate sulcus; A.S.: arcuate sulcus.

time-amplitude discriminators for online spike sorting. We used 0–4 single units for each electrode, the exact number varying from day to day, along with an optional multi-unit classification. Single units were preferentially included by signal-to-noise ranking. We collected data from 18 separate BCI experiments from monkey G, each experiment containing many hundreds of trials.

For monkey H, we used our second generation automated spike sorting infrastructure. This setup used the same methods as those for monkey G to determine the number of units and their overall characteristics. However, time-amplitude discriminators were not used for classification. Rather, we directly classified units within the same principal components space discussed above. We included multi-unit activity along with 0–5 single units per electrode. An additional ANOVA criteria was applied to include only units that were significantly modulated by reach target direction during the delay period ($p < 0.01$). We collected data from 40 separate BCI experiments from monkey H, each experiment containing many hundreds of trials. In BCI experiments with monkey H, different T_{skip} times (150–250 ms) were chosen on an experiment-by-experiment basis based on the cross-validated performance of the training trials, but the majority of experiments were conducted with T_{skip}=150 ms.

With the aforementioned selection criteria, ∼70–90 neural units were used in our highest performance BCI experiments with monkey H. To better understand what proportion of single units and multi-units were recorded, we examined neural data from our control experiments, fixing T_{skip} to 150 ms and T_{int} to 100 ms. For monkey H (using dataset `H20041217`; 8 targets), there were 25 tuned single units and 89 tuned multi-units (tuning assessed with ANOVA, $p < 0.05$). For monkey G (using dataset `G20040508`; 8 targets), there were 26 tuned single units and 65 tuned multi-units. We evaluated the sort quality of a unit (single versus multi) using all spiking data in each experiment and based on the discriminability between units in the modified principal components space (Sahani, 1999).

Though each data point in Fig. 2b of the main text represents performance consolidated over hundreds of trials in a given session, we would ideally replicate experimental conditions and repeat experiments over multiple sessions. Practically, the electrode array can only provide a quasi-stable number of neurons over a relatively short time (2–3 months). We chose instead to sample the fundamental design parameters (target configuration and T_{int}).

Models and decoding

The possible reach target locations were arranged in the patterns shown in Fig. 3 of the main text for monkey H. In all cases, we were careful to avoid placing targets directly below the center touch cue since this location would be obscured by the monkey's hand. Splaying targets out angularly, as opposed to linearly, recognizes the observation that most PMd neurons modulate their firing rate more for the upcoming reach direction than for the upcoming reach distance. We also explored two annular rings (8 and 16 target tasks) to demonstrate 2-dimensional target selection. Finally, we noticed that perturbing targets away from high-symmetry locations resulted in minor ITRC improvements. The performance improvement was partially due to the tuning properties of the particular neural units recorded from our electrode array. This suggests that true optimization of target placement based on neural response functions could increase the ITRC and requires future experiments.

Performance Calculations

We took a conservative approach in computing BCI performance. Specifically, we considered only sustained BCI trials. All BCI trials are not equivalent in their timing characteristics. In Fig. 1b of the main text, the first BCI trial contains a large center touch hold time. This period allows the monkey to reset its behavioral state after an immediately preceding reach trial. Consequently, the monkey is not being requested to rapidly switch his plan from a previous BCI trial. Including this particular trial's success or failure in our performance numbers is not a valid indication of *sustained* performance and could unduly inflate performance results. For the particular chain of trials shown in Fig. 1b, we only include trials #2 and #3 in our average performance results. As mentioned before, we take the whole trial time, consisting of T_{skip}, T_{int}, and $T_{\text{dec+rend}}$, when calculating all results that depend on the rate of target presentations.

Information transfer rate capacity calculation

Information theory has been used previously in neuroscience to estimate the information content in neural spike trains or other neural activity. This is not what we did; we did not attempt to estimate the intrinsic information content of the neural signals, at least not in any direct fashion. Instead, we calculate how much information, quantified in bits much like transmission of data over over a modem, can be extracted from the subject's thoughts, by way of our prosthetic system (which includes the entire signal path, from electrode recordings to target decoder).

We start by testing whether the target predicted from neural activity coincides with the target presented. If there is a match, the trial is deemed correct. Otherwise, the trial is an error trial. For error trials, we note the normalized frequency at which particular targets are decoded given a particular target presented. This allows us to characterize the "channel" of the communication system.

With these measurements, there are three ways in which to assess the information transfer (IT) per trial. For all calculations, the key quantity of interest is the mutual information metric:

$$I(X;Y) = H(X) - H(X|Y) = -\sum_{x} p(x) log_2 p(x) \; - \left(-\sum_{x,y} p(x,y) log_2 p(x|y) \right) \tag{1}$$

This equation simply states that the mutual information (I) between the set of presented targets (X) and estimated targets (Y) is the difference in entropy (or uncertainty) of the presented target set ($H(X)$) and the entropy after making an estimation ($H(X|Y)$). The experimental data is used to compute $p(y|x)$ which are the fractional occurrence of each estimated target y given a specific presented target x. The other quantities of interest are found from basic probability theory; $p(x,y) = p(x)p(y|x)$ and $p(x|y) = \frac{p(x,y)}{p(y)}$.

If Y provides a perfect estimate of X, $H(X|Y) = 0$; hence, the information transfer is maximal and equal $H(X)$, or the information contained in the presented stimuli. Taking an example of 8 targets, all presented with equal frequency, $H(X) = 3$, $p(x,y) = \frac{1}{8}$ if $y = x$ and 0 otherwise.

1. `Level-1 IT approximation` — convert the average prediction accuracy across an experiment to bits per trial. Again, if we have 8 targets, all presented with equal frequency, $p(x,y) = \frac{p_c}{8}$ if $y = x$ and $p(x,y) = \frac{1-p_c}{7\times 8}$ otherwise, where p_c is the fraction of occurrences that the estimated target matches the presented target averaged over *all* target presentations. (The number 7 appears in the denominator to equally distribute error across the remaining $y \neq x$ targets.)

2. `Level-2 IT approximation` — take into account error structure by computing the true mutual information between presented targets and decoded targets. This provides a more accurate representation of $p(x,y)$. In other words, every element of $p(x,y)$ is the fractional occurrence of the presented-estimated target pair (x,y), measured from our experiments. This can be a more accurate representation of information transfer. If, for example, errors are always distributed adjacent to the correct target, such a pattern is useful and taking it into account will lead to increased information transfer.

3. `Information Transfer Capacity (ITC)` — compute the full capacity of the communication system using the Blahut-Arimoto algorithm. The algorithm attempts to find the "capacity" (C) of the channel, namely the bits per use of the channel (averaged over many uses of the channel) such that there is zero probability of error.

$$C = \max_{p(x)} I(X;Y). \tag{2}$$

The algorithm starts with a guess for $p(x)$ and iteratively improves the estimate by solving successive constrained maximization problems with Lagrange multipliers until C converges to its global optimum (Cover & Thomas, 1990). Unlike the level-2 approximation, the system is not constrained to the relative frequencies of presented targets, $p(x)$, used during data collection. Importantly, this approach yields a system that utilizes certain targets (e.g., those that can be decoded more accurately) more often than other targets.

Supplementary Table S1: Comparing Methods for Calculating Information Transfer

	Targets		Performance			
	no. of targets	max bits/trial (max bpt)	accuracy (%)	level1-IT (bpt)	level2-IT (bpt)	ITC (bpt)
H	8	3	68.9%	1.2	1.6	1.9
	8	3	71.4%	1.3	1.6	1.7
	16	4	51.1%	1.1	1.9	2.2
G	4	3	93.0%	1.5	1.6	1.6
	8	4	73.5%	1.4	1.8	1.9
	8	4	76.8%	1.6	2.1	2.1

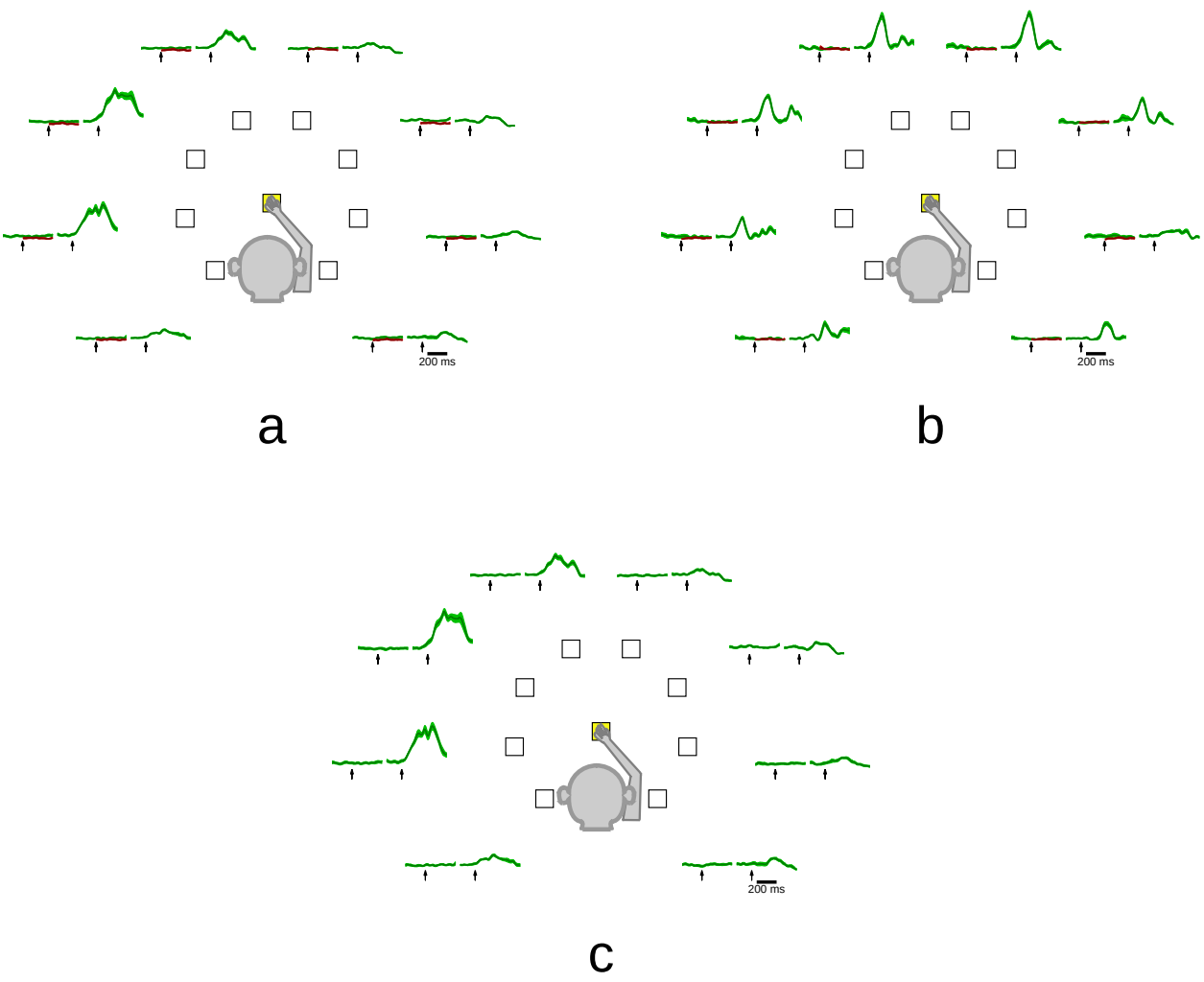
Table S1 shows the values obtained when using these different methods of calculating IT for a few representative BCI experiments. In general there was a large gain between the level-1 IT and level-2 IT calculations but only a modest gain ($\sim$15%) between the level-2 IT and ITC calculations. One could use any of these three methods of computing IT to then produce an information transfer rate (by dividing by the entire trial length, $T_{\text{skip}}+T_{\text{int}}+T_{\text{dec+rend}}$). The ITC is the standard performance metric in information theory (Shannon, 1948), and we use it to obtain the ITRC ($\frac{\text{ITC}}{\text{total trial length}}$) and to optimize our BCI.

8 EMG

We measured EMG from monkey G to verify that the neural activity was not a byproduct of minor limb movements during the delay period of our experiment. The aim was to ensure that the BCI system is operating with motor planning activity as opposed to movement execution activity. This is an important requirement if such a prosthetic system is intended for paralyzed patients. Supplementary Fig. S7 shows the data from three different muscles in monkey G. There is no noticeable difference in EMG activity between the periods before and after target presentation for real reach trials. Furthermore, there is no significant tuning in the EMG activity for target direction during the time period 50 to 300 ms after target presentation (ANOVA, $p >> 0.05$). We also measured EMG activity while presenting targets at a rapid pace ("rapid condition"), akin to the behavioral conditions present in BCI experiments.[4] Results were very similar to those of the real reach trials — there was no elevated activity after target presentation and there was no target-specific tuning in the EMG signal during the delay period.

The lack of tuned EMG signal in the delay period was typical across other monkeys in our laboratory. While we did not measure EMG from monkey H, the endpoint position of the monkey's fingertip did not move with respect to the target direction during the delay period. Our hand tracking apparatus has a sub-millimeter resolution.

[4]There was no movement epoch for the rapid condition, much like there was no such period for prosthetic cursor trials during BCI experiments.



Supplementary Figure S7: EMG measurements for monkey G plotted in arbitrary units. Data was collected for real reach trials (green) as well as trials with rapid presentation of targets (red). The first arrow designates the target presentation time and the second arrow marks 150 ms before the start of the movement. The delay period and movement periods are separated by a slight gap to allow for differing delay periods across trials. **a**. Measurements from the deltoid muscle. **b**. Measurements from the biceps muscle. **c**. Measurements from the triceps muscle. We only collected data from real reach trials for this muscle.

References

Carmena, J. M. *et al.* Learning to control a brain-machine interface for reaching and grasping by primates. *PLoS Biology* **1**, 193–208 (2003).

Cisek, P. & Kalaska, J. F. Simultaneous encoding of multiple potential reach directions in dorsal premotor cortex. *J. Neurophysiol.* **87**, 1149–1154 (2002).

Cover, T. & Thomas, J. *Elements of Information Theory* (John Wiley, New York, 1990).

Crammond, D. J. & Kalaska, J. F. Modulation of preparatory neuronal activity in dorsal premotor cortex due to stimulus-response compatibility. *J. Neurophysiol.* **71**, 1281–1284 (1995).

Crammond, D. J. & Kalaska, J. F. Prior information in motor and premotor cortex: Activity during the delay period and effect on pre-movement activity. *J. Neurophysiol.* **84**, 986–1005 (2000).

Kalaska, J. F. & Crammond, D. J. Deciding not to go: Neuronal correlates of response selection in a go/nogo task in primate premotor and parietal cortex. *Cerebral Cortex* **5**, 410–428 (1995).

Sahani, M. Latent variable models for neural data analysis. *Ph.D. Thesis, Computational and Neural Systems, California Institute of Technology* (1999).

Santhanam, G., Sahani, M., Ryu, S. I. & Shenoy, K. V. An extensible infrastructure for fully automated spike sorting during online experiments. *Proceedings of the 26th Annual International Conference of the IEEE EMBS, San Francisco, CA* 4380–4384 (2004).

Serruya, M. D., Hatsopoulos, N. G., Paninski, L., Fellows, M. R. & Donoghue, J. Instant neural control of a movement signal. *Nature* **416**, 141–142 (2002).

Shannon, C. E. A mathematical theory of communication. *Bell System Technical Journal* **27**, 379–423 and 623–656 (1948).

Shoham, S., Halgren, E., Maynard, E. M. & Normann, R. A. Motor-cortical activity in tetraplegics. *Nature* **413**, 793 (2001).

Taylor, D. M., Tillery, S. I. H. & Schwartz, A. B. Direct cortical control of 3D neuroprosthetic devices. *Science* **296**, 1829–1832 (2002).