Volitional control of neural activity: implications for brain–computer interfaces

Eberhard E. Fetz

Department of Physiology and Biophysics, Washington National Primate Research Center, University of Washington, Seattle, WA 98195-7290, USA

Successful operation of brain–computer interfaces (BCI) and brain–machine interfaces (BMI) depends significantly on the degree to which neural activity can be volitionally controlled. This paper reviews evidence for such volitional control in a variety of neural signals, with particular emphasis on the activity of cortical neurons. Some evidence comes from conventional experiments that reveal volitional modulation in neural activity related to behaviours, including real and imagined movements, cognitive imagery and shifts of attention. More direct evidence comes from studies on operant conditioning of neural activity using biofeedback, and from BCI/BMI studies in which neural activity controls cursors or peripheral devices. Limits in the degree of accuracy of control in the latter studies can be attributed to several possible factors. Some of these factors, particularly limited practice time, can be addressed with long-term implanted BCIs. Preliminary observations with implanted circuits implementing recurrent BCIs are summarized.

(Received 20 December 2006; accepted after revision 11 January 2007; first published online 18 January 2007)

Brain–computer interfaces (BCI) and brain–machine interfaces (BMI) convert neural activity at the level of neuronal action potentials, ECoG, or EEG into signals that control computer cursors or external devices. The BCI paradigm bypasses the normal biological pathways mediating volitional movements and employs upstream neural activity that may have a complex relationship to motor or cognitive behaviour. The transform between this neural activity and the required control parameters can be facilitated by sampling relevant activity in appropriate brain regions, such as motor cortex cells involved in limb movement. Conversion of these signals can be further aided by appropriate transform algorithms to generate the requisite control parameters. But even with the best matches and the optimal algorithms, accurate device control under diverse behavioural conditions depends significantly on the degree to which the neural activity can be volitionally modulated. Here we review evidence for such volitional control in a variety of neural signals, with particular emphasis on activity of single and multiple neurons. For further discussion of control of EEG, ECoG and field potentials, see other papers in this issue; Barber et al. (1971–1977); Wolpaw et al. (2002). The evidence for volitional control comes from conventional experiments that relate neural activity to behaviour, and emerges even more directly from studies using biofeedback and BCI.

Volitional activation associated with behaviour

The most obvious place to find cortical signals directly associated with volitional movements is primary motor cortex, where activity of accessible neurons is closely correlated with voluntary limb movement. Innumerable studies have demonstrated that cells in motor cortex and various premotor areas discharge with execution of voluntary movements in relatively specific and reliable ways. The diverse range of limb movements and the flexibility of digital control must clearly be correlated with correspondingly flexible activation of cortical cells that generate these movements. Relationships to movements can also be seen in cortical regions beyond traditional motor areas. In primary somatosensory cortex many cells that exhibit classic sensory responses to peripheral stimulation also fire prior to active movements, much like precentral motor cortex cells (Soso & Fetz, 1980); over half of the postcentral cells began discharging prior to activation of agonist muscles, revealing the existence of a central volitional drive that is superimposed on their peripheral input. Multiunit recordings in diverse cortical...
areas reveal that the parameters of free limb movements can be predicted from the activity of neurons in different pre- and postcentral cortical areas, with varying degrees of accuracy (Wessberg et al. 2000; Carmena et al. 2003).

Neurons in motor areas often fire also with imagined movements in the absence of execution. PET and fMRI studies have shown that many cortical areas associated with generating volitional movement are also activated when the subject simply imagines making the movement (Jeannerod, 1995; Roth et al. 1996; Jeannerod & Frak, 1999; Niyazov et al. 2005). Motor imagery is also effective in modulating synchrony and power in the EEG and ECoG (Pfurtscheller & Neuper, 1997; McFarland et al. 2000; Pfurtscheller et al. 2000; Leuthardt et al. 2004). Activation with motor imagery is further demonstrated by the decreased thresholds for evoking movements with transcranial magnetic stimulation (Kasai et al. 1997; Fadiga et al. 1999; Stinear & Byblow, 2003; Niyazov et al. 2005; Fourkas et al. 2006).

In addition to real or imagined movements, many cortical cells are modulated with movement preparation. This has been amply documented in studies that involve an instructed delay period, in which cortical cells may modulate their activity during the interval between the instructional cue and the ‘go’ signal (Wise et al. 1983; Kurata & Wise, 1988; Alexander & Crutcher, 1990; Riehle & Requin, 1995; Crutcher et al. 2004). Occurring after the end of any sensory response to the cue and well before the onset of the triggered movement, this instructed delay period activity may code information about the cue or preparation to move, but in either case reflects a volitionally generated activity. Neural activity associated with specific motor planning has been demonstrated in posterior parietal areas and may provide useful signals for decoding intended movements (Snyder et al. 2000; Shenoy et al. 2003; Musallam et al. 2004; Santhanam et al. 2006).

Neurons in sensory association areas are also volitionally activated in conjunction with cognitive imagery. In the temporal lobe many single neurons that respond selectively to a particular visual stimulus are in addition specifically activated during imagistic recall of the same effective stimulus (Kreiman et al. 2000). Thus, internal representations of stimuli and movements often employ many of the same neurons involved in overt sensory or motor behaviour. Beyond representations of sensory and motor events, internal cognitive activity like ‘thinking’ must also have neural correlates and these also represent volitionally controllable processes. These neural activities are independent of sensory input or motor output, and indeed operate autonomously because they are effectively buffered from peripheral activity.

Recent fMRI studies have shown that volitional shifts in attention activate widespread cortical areas in the absence of any sensory or motor correlates (Kastner et al. 1999). When subjects are fixating on a target spot and are cued to shift their attention to another part of the visual field, anterior cortical sites exhibit strong increases in activation, almost as large as the responses to an overt visual stimulus. Even primary visual cortex shows the effect of volitional shifts of attention, in the absence of any visual stimulus.

Thus, conventional experiments have revealed a range of circumstances in which central control of neural activity is evident. Volitional input could be considered to reflect an activating modality existing in addition to the better-studied sensory and motor modalities. The degree to which it is available for BCI/BMI control signals remains to be empirically determined. Conventional experiments, such as those described, are typically designed around a particular behaviour, and indirectly reveal the volitional components of correlated neural activity. Reversing this paradigm, biofeedback experiments directly elicit the volitional control of neural activity and allow the correlated behaviour to emerge.

**Volitional activation revealed by biofeedback**

The volitional drive on cortical neurons can be demonstrated directly by operantly training subjects to control the activity of neural activity with biofeedback. For example, operant conditioning experiments showed that monkeys were able to quickly increase and decrease the activity of motor cortex cells when rewarded for these changes (Fetz, 1969; Fetz & Baker, 1973). The degree to which cell activity met the criterion for reward was continuously represented in the displacement of a meter arm, whose rightward position corresponded to the threshold for the feeder discharge. Once the monkeys had discriminated this feedback they were able to drive the meter arm with newly isolated units and could modify their control strategy within minutes as the reward criteria were changed. Figure 1 shows an example of differential control of two neighbouring motor cortex cells. The firing rate of the unit with the larger action potential could be increased independently of the rate of the ‘smaller’ unit, and vice versa. Moreover, the monkey could also decrease the rate of the large unit (after several minutes of attempting increases, which had been previously rewarded). This bidirectional volitional control eliminates explanations involving non-specific effects like arousal or reward expectancy. Interestingly, these two units both responded reliably to passive extension of the knee, showing again that the central volitional drive on cells is controllable independently of peripheral input.

As might be expected, the operantly rewarded responses of many motor cortex cells were associated with active limb movements (Fetz & Baker, 1973). Indeed, the original rationale for these experiments was to identify the movements correlated with operant bursts of neurons as a motor analogue of sensory receptive fields. In many cases as the monkey continued to drive the rewarded unit,
the movements became more specific and often dropped out entirely. This dissociation has also been observed in studies in which cortical cell activity was used to drive a robotic arm or cursor, as described below (Chapin et al. 1999; Taylor et al. 2002; Carmena et al. 2003). Again, the ready dissociation between centrally driven activity and previously correlated movements speaks to the independence of the volitional drive on the cell from the motor circuits that generate active limb movements. It should also be noted that the monkeys activated some motor cortex cells for operant reward without ever making any observed movements (Fetz & Finocchio, 1975). Motor cortex neurons that were reliably associated with EMG activity in particular forelimb muscles could be readily dissociated from EMG when the rewarded pattern involved cell activity and muscle suppression (Fetz & Finocchio, 1975). This rapid dissociation of cell and muscle activity may reflect the rapid switching that is possible between imagining and executing movements. Given that the same cortical cells can be involved in both, and that central representations can be dissociated from action, the observed dissociation is readily explicable. Alternatively, the dissociation may be interpreted to demonstrate that cortical neurons have a lower recruitment threshold than motoneurons and that individual cortical neurons have a significant flexibility in being recruited during movement generation.

These studies are representative of a large body of experiments that have investigated the direct control of neural activity in the CNS through biofeedback (Barber et al. 1971–1977; Chase, 1974; Birbaumer & Kimmel, 1979). Given explicit visual feedback, subjects could volitionally control a number of physiological parameters that would otherwise remain unconscious. Volitional control of the activity of single neurons was initially investigated with single motoneurons through biofeedback of single motor unit activity (Harrison & Mortensen, 1962; Basmajian, 1963). Biofeedback worked well for activating low-threshold motor units in isolation, but not high threshold units; attempts to reverse recruitment order of motor units largely failed to demonstrate violations of the size principle. Olds pioneered CNS unit conditioning studies by operantly rewarding rats to increase the activity of midbrain neurons using intracranial stimulation (Olds, 1965). Biofeedback control of autonomic activity was also explored extensively, as described in Barber et al. (1971–1977) and Birbaumer & Cohen (2007).

---

![Figure 1. Operant conditioning of differential firing rates of two neighbouring motor cortex neurons](image_url)

Points plot 1 min average rates of large and small unit (L and S). ‘Operant level’ is activity prior to conditioning, with monkey seated in primate chair. Reinforcement periods are labelled by ‘↑’ and ‘↓’ indicating whether activity of the unit drove the biofeedback meter arm towards or away from level for triggering feeder. During ‘SΔ’ (time-out) periods feedback meter and feeder were turned off. (Used with permission from Fetz & Baker, 1973.)
Figure 2 illustrates the basic components of biofeedback experiments. The defining feature is the feedback about the state of the controlled variable made explicitly available to the ‘volitional controller’ – namely, the rest of the brain. The brain in turn uses the feedback to modify the controlled variable. In animal experiments additional feedback is often provided by rewarding the appropriate changes. An important concomitant of the reinforced activity is the correlated activity, which may have a causal relationship with the controlled variable or may be only adventitiously associated. For example, in biofeedback conditioning of single motor cortex cell activity, the correlated responses included the causally related activation of those cells directly driving the reinforced neuron, as well as associated motor activity that could be adventitiously related to the cell activity and be dissociable. Similarly, motor activity could affect many different conditioned variables – for example absence of movement enhances the precentral mu or beta rhythm (Pfurtscheller, 1981), motor activity is associated with hippocampal theta rhythms (Black, 1972), and closing the eyes enhances the appearance of occipital alpha activity (Mulholland & Eberlin, 1977; Ancoli & Kamiya, 1978). In many clinical applications of biofeedback the point of controlling the feedback variable (e.g. scalp temperature) was to change the correlated variable (blood flow and associated migraine headaches).

The black-box diagram in Fig. 2 is intended to identify relevant components, but of course separates these components artificially, since all are interacting parts of the volitional controller. Under certain circumstances additional relationships can exist. For example, the delivery of feedback or reward could itself have a direct effect on the reinforced activity. In such cases demonstrating bidirectional volitional changes in the reinforced activity would provide an important experimental control.

**Volitional activation revealed by BCI and BMI studies**

The volitional control of cortical cell activity has now been dramatically demonstrated in numerous BCI and BMI studies in which primates controlled the position of cursors or robotic arms with cortical activity under closed-loop conditions (Serruya et al. 2002; Taylor et al. 2002; Carmena et al. 2003). Under ‘open-loop’ conditions, the activity of neural populations could be linearly transformed to the 3-D coordinates of the monkeys’ hand as they retrieved food from a well and brought it to their mouth (Wessberg et al. 2000). Interestingly, the conversion parameters obtained for one set of trials provided increasingly poor predictions of future responses, indicating a source of drift over tens of minutes in the open-loop condition. This problem was alleviated when the monkeys observed the consequences of their neural activity in ‘real time’ and could optimize cell activity to achieve the desired goal under ‘closed-loop’ conditions. For example, monkeys could successfully acquire targets on a two-dimensional workspace (Serruya et al. 2002) or in virtual 3-D space (Taylor et al. 2002) with a cursor driven by activity of 10–30 motor cortex neurons. More recently, the weighted activity of cell ensembles recorded over many cortical areas was used to control a robotic arm to reach and grasp objects (Carmena et al. 2003). Significantly, several of these studies also demonstrated the ability to extract movement predictions from neurons in postcentral as well as precentral cortical areas (Wessberg et al. 2000; Carmena et al. 2003) (Fig. 3). Precentral motor cortex cells provided the most accurate predictions of force and displacement, but neurons from many other areas also
provided significant predictions. The prediction accuracy increased with the number of cells included, albeit with diminishing returns.

Human subjects could also exhibit cursor control with activity derived from an indwelling electrode (Kennedy et al. 2000) or from patterns of EEG activity (Wolpaw & McFarland, 2004). Most recently a paraplegic patient demonstrated significant control of a 2-D cursor and robotic arm with decoded activity of large populations of motor cortex neurons (Hochberg et al. 2006; Donoghue et al. 2007).

The basic BCI/BMI paradigm (Fig. 4) is essentially identical to the biofeedback paradigm. One emphasized difference is the transform algorithm converting neural activity to the control parameters needed to operate the device. This interposes an intermediate stage that may complicate the relationship between neural activity and the final output control of the device. The explicit reward loop has been eliminated to suggest that the volitional controller is typically motivated to operate the controlled device, although many animal experiments also employ a reward.

The relationship between the neural activity that is recorded and the correlated activity is again a significant issue. Many BMI studies first obtain an optimal basis for brain control by recording the neural activity associated with real or imagined limb movement and deriving appropriate transform algorithms (Chapin et al. 1999; Carmena et al. 2003).
Taylor et al. 2002; Carmena et al. 2003; Hochberg et al. 2006). When this algorithm is used to switch control of the device to neural activity, the previously associated movements can drop out with remarkable ease. Similarly, subjects may initially use some mental imagery to evoke ECoG activity that influences a cursor movement, but after a period of practice they often feel that they control the cursor more directly and drop the original mediating manoeuvre (Leuthardt et al. 2004). This flexibility of internal representations underlies the ability to cognitively incorporate external prosthetic devices into the body image, and explains the rapid conceptual adaptation to artificial environments, such as virtual reality or video games.

A comparable and related flexibility is demonstrated by the neural mechanisms that buffer mental activity from sensory input and motor output. Mental activity must be shielded from sensory disruption in order to operate independently of environmental events. It must also be dissociated from motor output to prevent imagined activity from being acted out and allow thinking to occur independently of movements. Yet these internal representations often employ many of the same neurons involved in overt sensory or motor behaviour. A highly flexible buffering component of mental operations allows central mechanisms to quickly switch between accessing sensory information or generating appropriate movements and performing the internal processing independently. These flexible switching operations are evident in BMI studies that tap the central activity and link it directly to external devices.

**Limitations on control for BCI and BMI**

Given the degree to which independent control of cortical units can be rapidly acquired in biofeedback experiments (e.g. Fig. 1), one might wonder why the control of BCIs and BMIs through neural activity is not more accurate than it is. Without minimizing the remarkable achievements of these studies, one can ask whether the limitations in accurate control are inherent or could be further addressed. There could be several possible explanations for these limitations. First, the complex transforms of neural activity to output parameters may complicate the degree to which neural control can be learned. In contrast to the relatively simple task of driving one or two cells in bursts while allowing free performance of any correlated responses, the requirement to modulate activity of a population to accurately control a transformed function may be more difficult because the effect of any particular cell is largely submerged in the population function. Moreover, activity of each cell in the population has some stochastic component which may conspire against learning optimal control of any particular cell (Carmena et al. 2003). Second, the degree of independent control of cells may be inherently constrained by ensemble interactions. A special example of such a constraint is the fixed relative recruitment order of motoneurons according to the size principle, which has foiled attempts to activate high threshold motor units independently of lower threshold units. Neural ensembles may have comparable limits on the degree to which individual elements can be independently activated. To the extent that internal representations depend on relationships between the activities of neurons in an ensemble, the processing of these representations involves corresponding constraints on the independence of those activities. These constraints may explain the diminishing returns obtained from increasing the number of neurons included in a linear filter (Carmena et al. 2003). The 'neuron dropping curves' representing the average accuracy as a function of the number of cells have extrapolated asymptotes below 100% for indefinitely large populations (Fig. 3). Yet, it remains possible that longer experience with the same neuronal ensembles could improve the achievable accuracy.

A third source of difficulty in achieving reliable control may come from employing adaptive decoding schemes. Although such adaptive algorithms are intended to automatically optimize control, they create a moving target for volitional modulation; the neural activity pattern that worked at one time may subsequently become less effective, requiring the learning of new patterns.

Finally, the ability to learn optimal control may be limited by the short and intermittent exposure times, dictated by the need to tether the subject to the requisite instrumentation. For example, a paraplegic subject that could practise neural control of a cursor only several hours a week demonstrated remarkable success in controlling a cursor movement, but nevertheless achieved a limited degree of accuracy (Hochberg et al. 2006). Intermittent sessions also involve possible changes in the recorded neuronal population, requiring the subject to relearn the task with a slightly different population of cells. These factors suggest that the range and reliability of neural control in BMI might increase significantly when prolonged stable recordings are achieved and the subject can practise under consistent conditions over extended periods of time. This would involve implantable circuitry that can monitor the same neural activity over many days.

**Implantable recurrent brain–computer interfaces**

Recognizing the need for implantable circuitry for further improvement in BMI control, many laboratories are developing compact, low-power integrated circuits (Mojarra et al. 2003; Obeid et al. 2004; Berger & Glanzmann, 2005; Mohseni et al. 2005). For example, we have investigated the operation of a small computer...
chip in conjunction with wire electrodes implanted in monkey motor cortex (Mavoori et al. 2005). This ‘Neurochip’ reliably recorded the activity of the same single neurons and two related arm muscles for weeks, storing raw and/or compressed data to memory for daily downloading via an infrared link (Jackson et al. 2007). The compact connections and self-contained circuitry makes unit recordings remarkably stable despite the unconstrained movements of the monkey in the cage. For many neurons the correlations between neural and muscle activity remained relatively stable, which bodes well for prosthetic applications.

The Neurochip can also operate in a recurrent loop mode, converting action potentials of a cortical neuron to stimuli delivered elsewhere in the motor system. Thus the cortical cell could directly control functional electrical stimulation of muscles, spinal cord or other brain regions (Jackson et al. 2006b). Continuous operation of such a recurrent BCI (R-BCI) should allow the subject to adapt to the artificial pathway and by appropriately modifying the neural activity, to incorporate its operation into normal behaviour. Such a R-BCI has obvious potential prosthetic applications in bridging lost biological connections, particularly when multiple parallel channels are implemented.

A second therapeutic potential is the possible strengthening of weak or impaired physiological connections. When the R-BCI was configured to connect neighbouring motor cortex sites, action potentials recorded at one site triggered synchronous stimulation at the second site (Jackson et al. 2006a). Continuous operation for a day or more of normal behaviour resulted in long-term changes in the output effects evoked from the recording site (Fig. 5). Surprisingly, these changes remained stable for over a week of testing after the conditioning paradigm had terminated. Such conditioning effects were not simply due to the stimulation alone, but involved time-dependent plasticity: testing numerous pairs of sites in this paradigm showed that none of the control sites exhibited any changes, and the

Figure 5. Continuous operation of a cortical recurrent BCI leads to long-lasting changes in physiological connections
Top: intracranial microstimulation at 3 different motor cortex sites with the monkey at rest evoked 3 different muscle responses (centre) and different isometric torques about the wrist (right). Arrows at right indicate means of 200 ms torque trajectories. Middle: conditioning involved 2 days of triggering microstimuli at site Nstim for every spike recorded at Nrec during free behaviour and sleep. Bottom: after conditioning the output effects evoked from site Nrec had changed to include those from Nstim, an effect that lasted beyond a week. A plausible mechanism is Hebbian strengthening of synaptic connections from Nrec to Nstim. (For further details see Jackson et al. 2006a.)
effect was obtained only when the delays between spikes and stimuli were less than 50 ms.

More sophisticated R-BCIs have been proposed for implementing recurrent computations in higher-order cognitive areas of the brain, like hippocampus (Berger et al. 2005). Conceivably, such neural prostheses might compensate for functions lost due to stroke or lesions by performing the lost computations and bridging the impaired regions. To operate as a ‘cognitive prosthesis’ the R-BCI would require effective communication between neural and electronic circuits at both input and output – a formidable technical challenge given the parallel distributed operations of biological neurons. In any case, technology is advancing rapidly and relentlessly, so we can anticipate further successes in developing continuously operating implanted BCIs.

References


**Acknowledgements**

This work was supported by the National Institutes of Health Grants NS12542 and RR00166.