Saving the baby: Toward a meaningful reincarnation of single-unit data

Eberhard E. Fetz

Department of Physiology & Biophysics and Regional Primate Research Center, University of Washington, Seattle, WA 98195
Electronic mail: *fetz@locke.hs.washington.edu*

Several commentators (**Alexander, Fuchs et al., Kalaska & Crammond, Lundberg, and Tanji**) described my target article as unduly pessimistic about the utility of single-unit recording in behaving animals. Alexander suggests I have concluded that "the response properties of central neurons are meaningless or uninterpretable." Having performed many chronic recording experiments, and having every intention of continuing such studies, I should first clarify that I think chronic unit recording is not only useful but quite essential for understanding neural mechanisms controlling movement. Without these experiments we would not be in the present situa-tion of having a wealth of information for evaluating their significance. My skepticism concerns more the utility of simplistic interpretations of neural response patterns in terms of correlated movement parameters and the as-sumption that neural networks operate through explicit coding of abstract conceptualized variables. In sensory systems the coding of stimulus features seems more secure, although **Bridgeman** raises corresponding con-

cerns even there; the idea that movement parameters are similarly coded in the motor areas is intuitively seductive but not convincingly supported by the accumulating experimental evidence. This does not mean, however, that chronic unit recording should be discarded along with the dubious interpretations. Although neural mechanisms controlling movement may not operate through explicit coding of movement parameters, the responses of neurons still offer crucial clues to neural computation.

The basic limitation of single-unit data is that they provide a very selective sample of a complex system, leaving a wide gap between highly restricted unit activity and overall system behavior. I would question the utility of bridging this gap by inference or leaps of faith. To better explain how neural networks control movement, we can fill this gap more effectively with neural network simulations. So, rather than being pessimistic about the utility of single-unit recording, I would be optimistic in hoping that single-unit data can finally be put into a causal framework: By incorporating the observed responses of units into neural network simulations we can generate more complete working models that can help explain the functional meaning of neural patterns.

The target article by **Alexander et al.** (this issue) seems to argue much the same point on the systems level that **Robinson**'s target article (this issue) and mine have argued on the cellular level, namely, that the operations of biological neurons do not conform to popular conceptual schemes that are imposed upon the recorded data. In his commentary **Alexander** suggests I would argue that "neural correlates of movement parameters . . . are only meaningful if they covary with the activity of the muscles with which that neuron is ultimately connected." Although I chose premotoneuronal (PreM) cells and their target muscles as useful examples of elements in the motor system with a correlational linkage that can be proven independently of their firing patterns, I do not believe that the relation of central neurons to movement parameters must all be analyzed in terms of these elements. In principle, higher-order cells could code parameters very well, and the literature is full of suggestive evidence for such coding, including, ironically enough, my target article. The tonic firing rate of certain PreM cells codes static muscle force much better than the motor units that produce that force, as shown in Figure 3 of my target article. Whereas individual motor units have a highly nonlinear relation to net force because of their recruitment threshold and saturation, many rubromotoneuronal (RM) and corticomotoneuronal (CM) cells have a linear relation to static force over a considerably wider range.

Alexander's commentary raises the important point that the evidence for the coding of a movement parameter can be strengthened when the relation between unit activity and a candidate parameter is tested under "a variety of system level input/output conditions." This strategy is crucial for determining whether a particular behavioral variable is reliably correlated with the activity of a neuron. For example, the contention that a central neuron codes direction of limb movement is strengthened when it correlates best with that direction independently of the force required to make the movement, or the part of extrapersonal space in which the movement is made. Unfortunately, there do not appear to be many motor cortical cells that pass such a battery of tests. As more behavioral situations are used to confirm the coding of a particular parameter, fewer neurons appear to correlate consistently with that parameter.

The use of multiple behavioral tasks is related to another excellent point raised by several commentators, namely, that additional techniques are important for providing a full picture of the role of cells in controlling movement. **Lundberg** and **Ioffe** mention the classical strategies of stimulation and lesions, which can be performed both electrically and pharmacologically. **Iansek** points out that the converging approaches should also include evidence from clinical observations. I agree that converging evidence from different techniques strengthens the functional arguments (Fetz 1981), but in the role of devil's advocate I would have to point out the pitfalls involved for interpreting unit data: Stimulation and lesions at best reveal the net effects of the majority of the affected cells and cannot reflect the function of every neuron in the affected region. Relying on the consequences of lesions and stimulation for interpreting the meaning of unit data has led to systematically underestimating the diversity of neural types in a region and the multiplicity of possible functions in which a particular region can participate.

Additional evidence from a number of procedures, including lesions, stimulation, and a repertoire of behaviors, would also be helpful in constructing realistic neural networks that could replicate the range of observations. The inability of biological neurons to consistently code a specific parameter under variable circumstances is bad news for simplistic concepts of neural coding, but these additional behavioral conditions are grist for the neural network mill. Simulating a battery of behavioral situations and responses to lesion and stimulation provides important additional constraints for neural networks. A neural network model becomes more plausible to the degree that it can simulate a range of experimental conditions.

In light of their curious antipathy to neural network modeling **Fuchs et al.** are relatively generous regarding my model and instead take issue with my comments about coding. They state that my target article describes paradoxical unit responses as presenting some "irreconcilable problems" for single-unit recording. In fact, I welcome those paradoxical responses as significant clues that mechanisms other than explicit coding are operating, and I agree entirely that these responses are "easily interpretable when there is a more complete understanding of the system"; this was precisely my point in discussing the insights provided by neural network simulations. We have two ways to deal with such paradoxical units: **Robinson**'s target article and the commentary of Fuchs et al. suggest that neural networks operate successfully despite the existence of so-called rogue cells because these cells are simply outnumbered by those with more appropriate activity. On the basis of simulations I would suggest the possibility that such rogue cells, which have counterintuitive properties individually, may have some rationale in the context of other such cells; that is, their inappropriate components may cancel, resulting in an appropriate contribution.

I would also take issue with the implication of **Fuchs et al.** that I consider nonlinear relations to be a problem.

Their statement that "the only disadvantage of nonlinear relations is that they tax the mathematical capabilities of the modelers" might apply to those theoreticians who try to capture the behavior of the networks in analytical form. In fact, most modelers are simulating networks with nonlinear units and have no problem with nonlinear behavior.

In comparing the oculomotor and somatomotor systems, this commentary indulges in the common conceit of some oculomotor physiologists that their system is somehow superior for being simpler. Thus it is claimed that "at least in the oculomotor system, a robust relation seems to exist several synapses from the motoneuron." Of course, one can find examples of such relations, if one is looking for them and is willing to ignore cells with more complex combinations of signals. However, not all oculomotor physiologists (e.g., **Robinson**) are convinced that the coding is as simple and clean and unequivocal as **Fuchs et al.** like to argue.

Gandevia points out that much of the complexity of neural activity in relation to movement may arise from the complexity of muscle activities involved in performing a motor task. This point is well taken and underlies the rationale for dealing first with much simpler alternating flexion/extension movements. As Gandevia further notes, even such simple movements may involve complicated coactivation patterns in proximal stabilizing muscles and distal finger muscles that can differ from reciprocal agonist/antagonist activity. This again is the rationale for focusing on those PreM cells that have a demonstrable correlational effect on the agonist muscles. Despite such restrictions, even PreM cells exhibit a variety of response patterns relative to their facilitated target muscles; one would think that inferring the relation of central cells to other muscles in more complex tasks would quickly become prohibitively complex.

Iansek raises the concern that if chronic unit recording data are presumed to be deficient, neural network models based on those data would also be flawed. In fact, the deficiency is not so much in the recorded data per se, which must be accepted as an experimental fact, but in the missing information about the rest of the system. Neural network simulations can help fill in the picture in a relatively objective manner, by tying the scattered observations together into a complete working model. So, neural network simulations actually provide a complementary method of analyzing and understanding the significance of the recorded single-unit data, even when the data alone provide a hopelessly selective sample of the system.

Kalaska & Crammond challenge the usefulness of obtaining a "complete description of the causal mechanisms for the planning and initiation of movement." This "Holy Grail" is dismissed as unattainable, which it is, if one considers the term "complete" to mean a comprehensive description of the state and connectivity of every relevant neuron. Obviously, such an exhaustive compilation of all the details in the biological nervous system is not only impossible to obtain but impossible to synthesize, and a description to this depth is unnecessary. Some simplifications must certainly be made, but we can still pursue solutions that are "complete" in the sense intended in my target article, namely, having a sufficient number of elements to implement a working dynamic solution. The network simulations are obviously simplified in many ways, but they can still provide a more complete representation of the neural mechanisms that can generate a behavior than a patchwork of scattered observations of individual neurons. Obtaining a causal model is ultimately a worthier scientific goal than intuitive reading of selected neural patterns.

Kalaska & Crammond appear to dismiss the neural network simulation in the target article as merely a "curve-fitting matrix" that transforms the inputs by low-pass filtering and weighted combinations of the inputs. Such networks, including this one, actually do more than simple linear operations on the inputs. As reviewed elsewhere (Fetz, in press), dynamic, recurrent neural networks can simulate the nonlinear operations of the oculomotor system (cf. **Robinson**'s target article, this issue), reflex responses with local sign (Lockery et al. 1990), autonomous oscillations that resemble central pattern generators (Rowat & Selverston 1991), and short-term memory tasks (Zipser 1991). My particular model was intended as an example of the method, rather than a physiologically realistic model of the sensorimotor system. Therefore, the fact that Kalaska & Crammond "see no evidence that it solves any of the sensory transformations required to convert target spatial location into a multidimensional intrinsic reference frame of muscular activity" is obviously because the network was never designed or trained with this task in mind. However, there is no fundamental limitation precluding the development of such a model.

Kalaska & Crammond make a good point that looking for parallels between activity patterns of hidden units and responses of biological neurons can involve the same sort of selection bias as correlating neural discharge to parameters of movement, and may be just as misleading. This concern can be addressed by determining whether the analogous neurons and hidden units each typify a representative set of elements involved in the task. In some cases the response properties are so distinctive that there is little question of the uncanny similarity of hidden-unit patterns (for example, the discharge patterns of cortical neurons and network units involved in short-term memory tasks; Zipser 1991). There is a very important difference between interpreting the meaning of response properties of single neurons in behaving animals and interpreting the response patterns of hidden units in the network. In the animal data, one is only guessing about their function. In the network models, one can demonstrate explicitly what that function is by tracing their connections. Therefore, to the degree that hidden-unit activity resembles activity of neurons in animals performing the same task, one can make inferences about the possible function of the biological patterns.

Kirkwood has a very particular bone to pick regarding the relation between postspike facilitation (PSF) and anatomical connections. We have taken PSF to indicate a correlational linkage between the PreM neuron and its facilitated target muscles, and have used it to define both operationally. But as Kirkwood points out, it is conceivable that some of these facilitations may be mediated not by a monosynaptic connection of the triggering cell, but by some other cells whose spikes are sufficiently synchronized with the trigger cell. As he indicates, we have analyzed this issue directly by cross-correlating CM and

neighboring cortical cells and found their cross-correlation peaks to be too broad to mediate the primary postspike effects (Fetz et al. 1991; Smith & Fetz 1989). Kirkwood suggests that more careful analysis of the PSF would be helpful in discriminating effects likely to be mediated by direct connections versus those mediated indirectly by synchrony, as previously described by Davies et al. (1985). This sort of analysis has been applied in a recent paper on the postspike effects of afferent fibers (Flament et al. 1992). A useful criterion is the latency of post*stimulus* effects, which can help define the minimal latency of postsynaptic effects of the trigger unit. Whether all the PSF in previous reports were mediated by direct monosynaptic connections or only some lesser proportion is not critical to the main point of my discussion, which concerns the response patterns of PreM cells. This issue would be relevant to my argument that PreM cells include a variety of response types if all examples of some particular response type were erroneously identified as PreM cells. This is highly unlikely, because each response type had many units with clear PSF that met the most stringent criteria.

The commentary by **Lundberg** makes several important points. Spinal cord neurons form a very important component of the motor system and we are indebted to Lundberg and his colleagues for their prodigious work in elucidating the segmental circuitry involved in reflex and voluntary movements. Although my target article describes primarily cortical and rubral PreM neurons, we do not assume that "only the monosynaptic pathways from the motor cortex and the red nucleus matter." Our experimental emphasis on the CM and RM cells is based primarily on their greater accessibility in behaving primates and on the fact that they do form a significant component of the supraspinal neurons that affect motoneurons directly. However, we do not believe that they are the only important controllers of motoneurons (Cheney et al. 1991). Regarding the points in the target article, they do provide a significant example of the sort of coding that can appear in cells directly linked to motoneurons. A major remaining experimental challenge is to elucidate the response properties of PreM cells in the spinal cord and we are currently trying to document cervical interneurons that affect motoneurons. The extensive work of Lundberg and colleagues (e.g., Alstermark et al. 1984; Baldissera et al. 1981; Illert et al. 1977; Jankowska & Lundberg 1981) will provide an essential context for identifying cells in such studies. Although the circuitry in the cat will provide important guidance in the monkey studies, significant differences between the cat and the primate, such as the existence of the monosynaptic corticomotoneuronal pathway, should also be remembered. The propriospinal neurons (PN) form an important disynaptic linkage from cortex to motoneurons in the cat (Illert et al. 1977), yet spike-triggered averages of EMG from cortex have so far failed to show that this linkage is sufficiently potent to mediate PSF. Determining whether similar PNs exist to the same extent in the monkey is a primary goal of our current investigations. We agree with Lundberg's suggestion that the CM cells may resemble the C3-C4 PNs in having two subgroups: those cells that project only to synergistic motoneurons and others that project both to motoneurons and Ia inhibitory interneurons (Fetz et al. 1990; Kasser & Cheney 1985).

Another difference between primate and cat may be the degree to which the lateral and medial descending systems play separable roles. **Lundberg** says it is noteworthy that complete transsection of the corticospinal tract and the rubrospinal tract in cats does not abolish commands for target-reaching and manipulatory movements. This result differs from those reported by Lawrence and Kuypers (1968) for the monkey, again suggesting significant interspecies differences. As Lundberg points out, reticulospinal pathways may also play a significant role in controlling primate motoneurons, another challenge for the spike-triggered averaging technique.

The relation of CM cells to movement is further discussed by **Lemon,** whose work has elucidated the function and connections of CM cells related to distal hand and finger muscles in natural precision grip movements. His work provides new insights into the relations between PSF patterns and responses of the monkey hand. Lemon raises the important point that it is helpful to document the response of PreM cells in relation to relatively normal limb movements in addition to a simple flexion-extension task. He says that when activity of CM cells is related to free hand movements "there is a congruency in the pattern of synaptic connectivity [shown by the PSF] and of recruitment during movement." An important caveat here follows from the fact that PSF can be detected only in those muscles coactivated with the cell; thus, during free movements involving variable activation of different muscles, this condition itself will tend to produce a congruence between the cell's facilitated muscles and the coactivated muscles.

Tanji raises an objection that calls for clarification. According to Tanji, my target article suggests "that no area-specific differences in properties of neuronal activity have been reported" and implies that "only neurons of the same type have been found everywhere in motor areas in more or less the same degree." In fact, my paper does state, perhaps not emphatically enough, that different cortical regions clearly show differences in the relative proportions of cells involved in different functions. Indeed, the work of Tanji and colleagues is particularly exemplary in providing persuasive evidence for regional specializations (e.g., Mushiake et al. 1991). Their work has documented the proportions of cells in different areas under behavioral tasks designed to elucidate these functional differences, and has provided ample data supporting specializations. My point is rather that cells of the same response type can be found distributed over many regions and that these like-minded cells probably form functional groups. The degree to which any particular cell type is found experimentally can also be proportional to the degree to which it is sought; therefore, recording bias should also be carefully controlled in experiments involving a search for different types of neurons. I agree entirely with Tanji's point that the experimental data "point to the presence of specialization in cortical areas" and with **Grobstein**'s similar point that we are "not in general dealing with a fully distributed system but rather with one having discrete 'information processing blocks.'"

What may have led to this misunderstanding is my statement that given a sufficient variety of cell types one can find examples to support any hypothesis from a completely random data set. This statement was designed to make a point but does not reflect a belief that the neural

data are in fact random. There is good evidence for preferential relations between neural discharges and components of movement, and efforts to resolve what those discharge patterns actually mean are certainly worthwhile. My main point is to caution against inferring their meaning by conceptual projections, as opposed to determining their meaning in a causal framework; the latter can now be approached by neural network simulations that replicate these patterns and provide a mechanistic basis for interpreting their computational significance.

References

Alexander, G.E. DeLong, M.R. & Crutcher, M.D. (1992) Do cortical and basal ganglionic motor areas use "motor programs" to control movemement? *Brain and Behavioral Sciences* 15:656-665.

Alstermark, B., Lundberg, A. & Sasaki, S. (1984) Integration in descending motor pathways controlling the forelimb in the cat. 11. Inhibitory pathways from higher motor centers and forelimb afferents to the C3-C4 propriospinal neurones. *Experimental Brain Research*, 56: 293-307.

Baldissera, F., Hultborn, H. & Illert, M. (1981) Integration in spinal neuronal systems. In: *Handbook of Physiology—The Nervous System II.* Brookhart, J.M. et al. (eds) American Physiological Society, Bethesda MD pp. 509-595.

Cheney, P.D., Fetz, E.E. & Mewes, K. (1991) Neural mechanisms underlying corticospinal and rubrospinal control of limb movements. *Progress in Brain Research, 87*: 213-252, .

Davies, J.G. McF., Kirkwood, P.A. & Sears, T.A. (1985) The detection of monosynaptic connexions from inspiratory bulbospinal neurones to inspiratory motoneurones in the cat. *Journal of Physiology* 368: 33-62.

Fetz, E.E. (1981) Neuronal activity associated with conditioned limb movements. In: *Handbook of Behavioral Neurobiology,* Vol. II: Motor Coordination, A.L. Towe and E.S. Luschei, eds., Plenum Press, pp. 493-526.

Fetz, E.E., Cheney, P.D., Mewes, K. & Palmer, S. (1990) Control of forelimb muscle activity by populations of corticomotoneuronal and rubromotoneuronal cells. *Progress in Brain Research 80*: 437-449.

Fetz, E.E. Toyama. K. & Smith, W. (1991) Synaptic interactions between cortical neurons. In: *Cerebral Cortex, Vol IX Altered Cortical States,* A. Peters & E. G. Jones, eds. Plenum Press, New York, 1-47.

Fetz, E.E. (in press) Dynamic Neural Network Models of Sensorimotor Behavior. In: *The Neurobiology of Neural Networks,* Daniel Gardner, Ed., MIT Press, Cambridge.

Flament, D., Fortier, P.A. & Fetz, E.E., (1992) Response patterns and post-spike effects of peripheral afferents in dorsal root ganglia of behaving monkeys. *Journal of Neurophysiology* 67: 875-889, 1992.

Illert, M. Lundberg A. & Tanaka, R. (1977) Integration in descending motor pathways controlling the forelimb in the cat. 3. Convergence on propriospinal neurones transmitting disynaptic excitation from the corticospinal tract and other descending tracts. *Experimental Brain Research* 29: 323-346.

Jankowska, E. & Lundberg, A. (1981) Intemeurones in the spinal cord *Trends in Neurosciences* 4: 230-233.

Kasser R.J. & Cheney P.D. (1985) Characteristics of corticomotoneuronal postspike facilitation and reciprocal suppression of EMG activity in the monkey. *Journal of Neurophysiology* 53: 959 - 978.

Lawrence, D.G. & Kuypers, H.G.J.M. (1968) The functional organization of the motor system in the monkey. II. The effects of lesions of the descending brainstem pathways. *Brain* 91: 15-36.

Lockery, S.R., Fang, Y. & Sejnowski, T.J. (1990) A dynamical neural network model of sensorimotor transformations in the leech. *Neural Computation* 2: 274-282.

Mushiake, H., Inase, M. & Tanji, J. (1991) Neuronal activity in the primate premotor, supplementary, and precentral motor cortex during visually guided and internally determined sequential movements. *Journal of Neurophysiology 66*: 705 -718.

Robinson, D.A. (1992) Implications of neural networks for how we think about brain function. *Brain and Behavioral Sciences,* 15:644-655.

Rowat, P.F. and Selverston, A.I. (1991). Learning algorithms for oscillatory networks with gap junctions and membrane currents. *Network* 2: 17-41.

Smith, W.S. and Fetz, E.E. (1989) Effects of synchrony between corticomotoneuronal cells on post-spike facilitation of muscles and motor units. *Neuroscience Letters,* 96: 76-81.

Zipser, D. (1991) Recurrent network model of the neural mechanism of short-term active memory. *Neural Computation* 3: 179-193.