Servo Action in Human Voluntary Movement

C. D. MARSDEN*, P. A. MERTON† & H. B. MORTON
National Hospital, Queen Square, London WC1

Muscular movements are under the control of a servo similar in many ways to those used in engineering control systems but with a subtlety of design not found in man-made servo mechanisms—including an automatic gain compensation for altered load which must be useful, for example, for adjusting to reduced *g* on the Moon.

IF in the course of a voluntary movement we meet with an unexpected obstruction we soon exert our muscles harder to overcome it. The question is, how soon? If an ordinary voluntary reaction time is involved the delay is not likely to be much under a seventh of a second (140 ms). On the servo theory of muscle control, however, the delay ought to be much less. This theory[1,2] supposes that the contraction of the main muscle is driven by contraction of specialized sensory structures in the muscle, the muscle spindles. The sensory endings on the spindles function as misalignment detectors and automatically turn on more contraction, via the stretch reflex arc, if the rate of shortening of the main muscle falls behind the shortening of the spindles. This type of follow-up servo action, akin to power-assisted steering on a motor car, would tend to make the main muscle shorten at the same rate as the spindles. If an obstruction were met with, the delay in turning on more contraction to overcome it would be only the brief delay associated with the operation of the stretch reflex.

So far the evidence for the servo theory from animal experiments has been equivocal[3,4], and there has been no convincing evidence for its operation in the human subject. In the present experiments, direct evidence has been obtained by interfering unexpectedly with flexion movements of the top joint of the human thumb.

This movement was chosen because it involves only one muscle, the flexor pollicis longus, which lies in the forearm. The proximal phalanx was held in a clamp so that the axis of rotation of the top joint coincided with the axis of a low-inertia electric motor (Printed Motors Limited, Type G9M4, used as a torque motor moving a fraction of a turn only). The spindle of this motor carried at one end a potentiometer, to record thumb position, and, at the other, an arm against which the pad of the thumb pressed to move or be moved by the motor. This arm incorporated a strain gauge for measuring the force exerted by the thumb. All the phenomena to be described can be seen on force records, but they are more elegantly demonstrated by using, as an index of the activity of the flexor muscle, its action potentials (electromyogram) led off by a pair of surface electrodes over the belly of the muscle in the forearm and then amplified and integrated. The slope of the integrated record gives the degree of contraction of the muscle.

There was a standing current in the motor which held the motor back against a stop and thus defined the starting position of the thumb. The subject made flexion movements in his own time, about once in 5 s. To standardize the movements he was given a tracking task on a cathode ray tube. By keeping two spots together he executed a 20° movement in 1.2 s. To smooth out irregularities several records were averaged in a 'Biomac' (Data Laboratories Limited).

All movements were begun against the constant force offered by the standing current in the motor, but about half way through perturbations were sometimes introduced. The sweep of the 'Biomac' was triggered at approximately 8° of flexion, and 50 ms later one of four things happened unpredictably, as determined by the output of a randomizing device: either the motor current did not change (a "control" trial); or it dropped rapidly to a fraction of its previous value and remained there for the rest of the movement, allowing the thumb to accelerate ("release"); or it rapidly increased to roughly double, applying a step of extra force to the thumb and driving it back ("stretch"); or finally (by feedback from the potentiometer on the spindle) the motor from then on was caused to simulate a spring, so that further movement was met with an increasing opposing force. The rate of the spring was such as to double the force in 3°. When it met the spring the thumb was almost arrested (a "halt" trial), but was not driven back.

The 'Biomac' sweep lasted 250 ms. From 50 ms to the end of the 'Biomac' sweep the subject's spot on the tracking cathode ray tube was extinguished, so that he had no visual clue as to what was being done to his thumb. In an experiment, the subject went on making trial movements until sixteen (or sometimes eight) of each kind, as they turned up at random, had been accumulated for averaging in the appropriate 'Biomac' channel.

The type of result obtained with the two authors (C. D. M. and P. A. M.) who acted as subjects is shown in Fig. 1*a*. The lower part gives the records of mechanical displacement and shows the average for the "control" trials, with the "release", "stretch" and "halt" records diverging from it after 50 ms. The electrical records above show that after a delay of about 60 ms the muscle attempts to compensate for these deviations from the desired track by reacting in the opposite sense. In the other subject, C. D. M., whose arm length is 12% less, the figure is more like 50 ms. These latencies are only a fraction of the shortest recognized voluntary reaction times, and they were the same in earlier experiments[5] in which the subject was faced with two rather than four choices at the 50 ms mark. Voluntary reaction times increase rapidly with the number of choices (Hick's law[6]); but here the latency is independent of the number of choices.

It seems safe to conclude that the response to stretch is reflex and we proceed on the assumption (later to be questioned) that it corresponds to the muscle-spindle-excited stretch reflex of decerebrate experimental animals. Our latencies for the stretch reflex agree sufficiently with those found in the human elbow flexor by Hammond[7]. As far as we are aware there are no figures available for the latency of the stretch reflex, as opposed to the tendon jerk, in animals. The fact that the halted and release responses have the same latency as the stretch reflex makes it likely that they are manifestations of the same mechanism. If this is accepted, then the release response is the negative of the stretch reflex and it demonstrates that the muscle must have been receiving excitation *via* the stretch reflex arc at the moment of release. (This argument is essentially the same as the original argument from the silent period

* Permanent address: Institute of Psychiatry, de Crespigny Park, London SE5.

† Permanent address: Physiological Laboratory, Downing Street, Cambridge.

which led to the servo theory[1].) The response to halting is, in effect, a stretch reflex without muscle stretch. It is the best piece of evidence we have for the original servo theory because, from the animal evidence, it is difficult to see how it could arise except by contraction of the spindles stretching the sense endings on them. And a surprisingly brisk contraction it must be, for (in this experiment, but not always) the slope of the "halt" response approaches that of the "stretch" response. This is even more conspicuous in Fig. 2*a*, to which we shall come later.

The sensitivity, too, of the system as revealed in the "halt" trials is remarkable. From the geometry of the tendon and thumb joint, the rate of tracking, the minimum time that must be lost in conduction in the reflex arc *via* the spinal cord and the latency of the servo response, it can be calculated that the misalignment signal calling for more contraction must have left the muscle when its degree of shortening was only some 50 μm (0.05 mm) less than it would have been if the muscle had gone on shortening unimpeded. And this figure is an upper limit which assumes that movement of the thumb stops at once when the extra resistance begins to develop, that the muscle does not continue to shorten against the various compliances in series with it, and that the stretch reflex is spinal. In the release trials, the misalignments which turn the muscle completely off must be of the same order of magnitude. The detection, within milliseconds, of these minute misalignments in an actively shortening muscle many centimetres long represents a considerable feat of sensory discrimination and we begin to see why muscles require the profuse and elaborate sensory apparatus they have long been known to possess and which has always hitherto been something of a mystery.

Returning to the consideration of Fig. 1*a*, as already noted, the response to stretch is not very much steeper than the response to halting, although, particularly in the early stages, the tracking error and, hence, the misalignment presented to the muscle spindles is several times greater during stretch. This

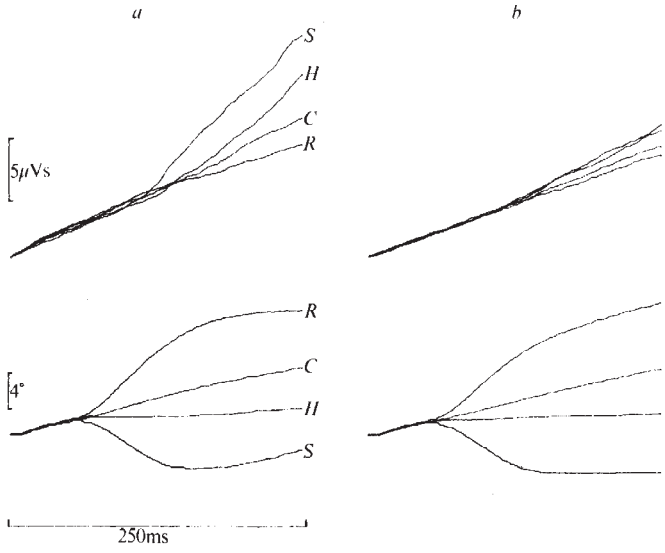

Fig. 1 Servo-type responses from the muscle flexing the top joint of the thumb. During the course of a tracking movement against an initially constant resistance, as described in the text, the movement may either be reversed (*S*) eliciting a stretch reflex, or halted (*H*), or allowed to accelerate by a reduction of the opposing resistance (*R*). In the control trials (*C*) none of these things happen. The initial resistance was a torque of 150 g cm (about 2% of the maximum torque that the subject could exert). The top records are the integrated electromyogram of the flexor muscle, the slope of the trace giving the degree of activation of the muscle. The bottom records give the angular position of the terminal phalanx of the thumb. Each trace is the average of sixteen trials, formed in a 'Biomac' averaging computer. *a*, Records taken with a normal thumb; *b*, records taken in the same experiment after anaesthesia of the thumb induced by occluding the blood supply for 1.5 h with a cuff at the wrist. Subject P. A. M.

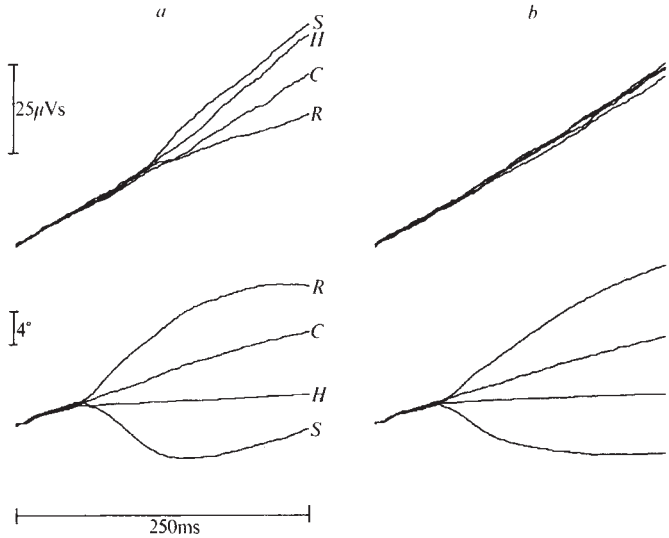

Fig. 2 As Fig. 1, but with the movements against a ten times greater initial resistance. Subject P. A. M.; an experiment on a different day from Fig. 1.

suggests that the spindle-based excitatory machinery is beginning to saturate, although the degree of contraction reached in these experiments, even in the stretch responses, can easily be shown to be only a small fraction (roughly 5%) of what the muscle can achieve.

The release trials, on the other hand, tend to emphasize the importance of the servo machinery: other experiments, in which action potentials were not integrated, showed that release could completely silence the muscle, with the usual brief latency. Merton's experiments[8] on the "silent period" during steady contractions in the human subject indicate that this is likely to be true at all levels of contraction up to maximal, suggesting that the saturation level of the servo machinery must be raised as the level of contraction rises, presumably by something akin to a change of loop gain.

This is not the only possibility. The theory of α–γ coactivation[9], which, in various forms, has been widely favoured[3,4,10], proposed that some excitation reached the muscle *via* the servo loop (the γ route) while some by-passed the servo loop and played directly on the main motoneurones (the α route). This theory has recently been greatly strengthened by the important direct human evidence of Vallbo[11]. One way in which large contractions could be obtained on this theory would be by increasing the α component after the γ route begins to saturate, without changing the loop gain. But in this case it should be more difficult with large contractions to silence the muscle by release, and the extra contraction turned on by stretch or halting should become a smaller proportion of the total.

To decide between these two mechanisms, experiments were performed in which the initial resistance offered to movement was increased ten-fold, with alterations in the forces during release and stretch so as to produce approximately the same displacements as before. (An initial resistance of this value requires roughly one quarter of the maximum force the long flexor can exert to overcome it.) The results of these experiments were unequivocal. At the high force the electrical activity of the muscle was much increased and the recording gain had to be reduced. In both subjects the shape of the electrical responses remained remarkably similar to those seen at the lower level of force. A typical record is shown in Fig. 2*a*. The response of the muscle to release, stretch or halt is given by the change in slope of the integrated electromyogram, so the general similarity of Figs. 1*a* and 2*a* shows that the *proportional* change in the degree of activation of the muscle in the release, stretch and particularly the halt trials remains effectively constant over this ten-fold change in initial force. Hence the muscle activation turned on by a given misalignment (the gain) is proportional to the initial force.

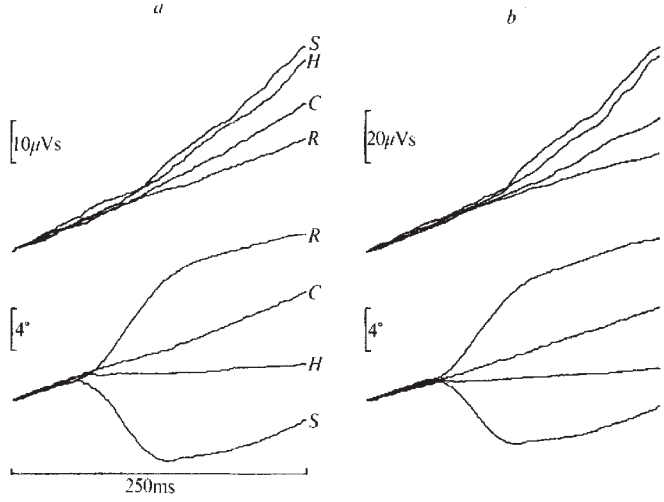

Fig. 3 Gain compensation during fatigue. The initial force in this experiment was low, the same as for Fig. 1. *a*, Fresh muscle. The muscle was then fatigued by applying maximal shocks to the median nerve at the elbow at 50 s^{-1} for 60 s. Recovery was prevented by a blood pressure cuff on the upper arm, arresting the circulation. *b*, Muscle fatigued. The electrical responses are of the same form as before fatigue, but scaled up by a factor of roughly 2, as shown by the calibrations. Force records (not illustrated) showed that the forces developed in the halt and in the other responses were the same as before fatigue. In these experiments each trace is the average of eight trials. The period of circulatory arrest was too short to cause ischaemic nerve block. Subject P. A. M.

According to Professor P. H. Hammond, whom we have consulted as a practising control engineer, change of gain with load is not a feature of man-made servo systems. Evolutionary pressures have produced a more flexible device than have engineering requirements. If an engineer designed an artificial arm, a powered prosthesis, with which one could wind up one's car to start the engine, incorporating conventional servo assistance to get over the compression strokes, and then used it to wind up a clock instead, it would break the mainspring when it came to the end !

One possible mechanism for the gain control in the servo loop postulates an input to the servo which, as the demanded force alters, has the function of altering the number of α motoneurones that are sensitive to spindle excitation, in proportion. A demand for increased force would be met (in whole or in part) by increasing the number of motoneurones responding to spindle input in proportion (or at a slower rate if the larger motor units are recruited later). Loop gain would automatically increase in proportion, as the number of active motoneurones increased. It is known that the number of α motoneurones active does increase with force and that all α motoneurones receive a spindle input. We propose that the spindle input does not lead to excitation of an α motoneurone unless the gain input is active at the same time—and *vice versa*.

To illustrate this proposal, imagine a motor car in which the mean speed for a journey was determined by altering the number of cylinders (representing motoneurones) in use. To keep the speed constant during the journey the accelerator could be depressed to climb a hill, or released when descending. The briskness of response to the accelerator (the gain available) would be proportional to the number of cylinders in use, and hence to the initial speed.

An important attraction of the idea of an input to the servo that causes a motoneurone to become accessible (or to become more accessible) to spindle excitation is that it would provide a "focusing" effect, permitting the higher centres to use servo action for discrete contractions of one small part of a large muscle, despite the known fact that excitation from a single spindle is often very widely distributed among the motoneurones of a large composite muscle, or to the motoneurones of more than one muscle. The need for such a "focusing" action in movements of individual fingers using the common flexor and extensor muscles in the forearm has been emphasized by Phillips[4].

At one time, for reasons which will emerge presently, we entertained the notion that signals from pressure receptors in the thumb itself were responsible for sensing the force exerted and turning up the gain appropriately[12], but this theory was made implausible by the behaviour of the servo during muscular fatigue. The muscle was fatigued by working it while the circulation was occluded by a blood pressure cuff on the upper arm. In the fatigued state the contractility of the muscle is lowered and it has to be more strongly activated to achieve the same force. This shows up as an increase in the electrical activity, which, as with increased load, entails a reduction in the recording gain. The force and displacement records were little altered as contractility fell, but the electrical responses were scaled up in the same way that they are scaled up when the demanded force is raised (Fig. 3). It seems therefore that muscular fatigue is compensated for by an increase of gain elsewhere in the servo of the same type as occurs with increased load. Because in the fatigue experiment the pressures on the thumb are unchanged, cutaneous receptors cannot be controlling the mechanism that boosts the gain.

Automatic compensation for load and fatigue have always been numbered among the supposed advantages of the muscle servo but the original theory[1] did not predict that the compensation would take the form or show the effectiveness that it does. Even at that time, however, it was noted that, contrary to expectation, when skilled finger actions, such as writing or playing musical instruments, are carried out during circulatory arrest, there is little falling off in performance almost until the muscles become so weak that they are unable to execute the necessary movements at all[13]. This now becomes intelligible.

Early on in these experiments we made a puzzling discovery which does not seem to have any correlate in the animal literature and which may cast light from an entirely new direction on the whole scene. If the thumb is anaesthetized, without interfering in any way with the muscle itself in the forearm, servo action is lost and the stretch reflex depressed[5]. The thumb may be anaesthetized either by a ring block with 'Xylocaine', or by a rubber band round the base of the proximal phalanx left on for 1.5 h or more, or by inflating a narrow cuff around the wrist for a similar period. We have done numerous experiments of this type on both subjects at low forces, always with a similar result (Fig. 1*b*); and we have confirmed that the same occurs in a fatigued muscle and at high forces (Fig. 2*b*). All the mechanical records are, in their first half, closely similar to those made before anaesthesia, but the responses in the electrical record are all greatly depressed throughout. Towards the end of the record, the mechanical responses in both Fig. 1*b* and Fig. 2*b* demonstrate the effects of the lack of servo compensation, by their wider or more persistent divergence from the control record.

An unexpected subjective observation of interest is that both subjects report that it needs more conscious effort to start the movement when the thumb is anaesthetic, particularly to restart it after it has been halted, and especially with the first few movements of a freshly anaesthetized thumb. Thus anaesthesia of the thumb provides both objective and subjective evidence of the advantage to the subject of servo action in his muscles.

Our experience is that anaesthesia depresses all the responses simultaneously at both high and low forces. Apart from providing further confirmation for the view that the stretch, halt and release responses share a common mechanism, this suggests that anaesthesia acts by turning down the gain of the servo loop, for if, on the alternative view, anaesthesia put the servo out of action by stopping spindle contraction, one would expect the release response to be lost first, followed by the halt response, the stretch reflex disappearing last. Whichever is the case, movements of an anaesthetic thumb must be produced by

the α route. The records in Figs. 1*b* and 2*b* indeed display in an extreme form the characteristics described above (in considering α–γ coactivation) for movements against a large resistance when the extra force was to be produced by increasing the α component without turning up the γ loop gain, that is, a reduction in that part of the contraction which is sensitive to stretch, halt, or release. (From another point of view the results can be regarded as establishing in the human subject, albeit with an anaesthetic thumb, the original thesis of Granit, Holmgren and Merton[9] that the same movement can use the α and γ routes in varying proportions.)

It might be held that all the responses in Figs. 1*a* and 2*a* are directly due to receptors in the pad of the thumb sensitive to the force to which it is subjected. But the fatigue experiments suggest that such receptors are unlikely to be involved in the gain control mechanism, and we have evidence that they are not necessary for the responses to stretch or halt, because normal responses of this kind are obtained when the thumb moves the motor by pulling on a string cemented to the thumbnail. The phenomenon of the silent period[8], in which a rise of force is associated with a complete cessation of muscle activation, is also evidence against this possibility. In animals, of course, stretch reflexes are obtained in isolated muscles with the rest of the limb denervated.

What, then, is to be made of this effect of peripheral anaesthesia ? The evidence cited earlier has shown that increasing activation of the servo, whether to overcome an extra load or offset fatigue, automatically boosts the gain (possibly by rendering more motoneurones susceptible to spindle excitation), and as a consequence servo action is focused on the muscle needed for the movement. We now learn that sensation in the thumb itself is necessary for that part of servo activation which turns up gain and focuses spindle excitation, but not, apparently, for that part which causes the spindles to contract. This power of thumb sensibility over the gain and focus control might be exerted either at spinal or at cerebral level. The latter possibility seems the more attractive. It could mean, for example, that the cerebral cortex cannot throw into play a mechanism to focus spindle excitation on to the thumb flexor if, because it is anaesthetic and not under visual control, attention cannot be focused on the thumb.

This is the time to recall Phillips's suggestion[4] that the stretch reflex may not (as hitherto implied here) be purely spinal but may involve what he calls a transcortical loop. In the baboon, and presumably in man, signals from muscle spindles reach the cerebral cortex[14], so that our conditional focusing may be an action on cortical rather than spinal motoneurones. A transcortical stretch reflex has not been clearly established in animals or in healthy human subjects, but something of the kind, with a latency similar to the latency we have found for the stretch reflex in healthy subjects, is seen in rare cases of myoclonic epilepsy. In Carmichael's case[15], investigated electrophysiologically by Dawson[16], myoclonic jerks could be elicited by stretch of muscle and only by that variety of mechanical stimulus. The pathological findings in this case were predominantly cerebellar degeneration (J. G. Greenfield, unpublished post-mortem report, the National Hospital, Queen Square). In this patient electrical stimulation of the ulnar nerve at the wrist evoked a large potential wave in the contralateral motor area 3 cm in front of the central sulcus (the position of which was verified at post-mortem) and, therefore, clearly in front of the area in which evoked responses to sensory stimulation are found in healthy subjects. It was followed by a generalized myoclonic jerk, but the latency of this jerk in the arm muscles was not recorded. Such measurements have, however, been made on other myoclonics by Fullerton and Giblin (reported by Halliday[17]), by Lance and Adams[18] and by Kugelberg and Widén[19]. In all these cases the latency of the myoclonic jerk was roughly double the latency of the evoked cortical wave. This would give a jerk latency of some 50 ms for the forearm muscles, agreeing with our healthy stretch reflex latency. The cortical origin of these myoclonic jerks was perhaps established most definitely in Kugelberg's and Widén's case[19], in which they were limited to one leg and were cured by excision of a small region of the leg area of the opposite motor cortex.

Transcortical stretch reflexes are thus very much on the cards. Jumping ahead of the evidence for the sake of a definite proposal to serve as a basis for further experiment, suppose that in a voluntary movement of the thumb cortical motoneurones connecting to both α and γ spinal motoneurones are activated. If tactile sensibility in the thumb is active, signals from muscle spindles in the thumb flexor arriving back at the cerebral cortex are given access to the α cortical motoneurones, thus completing a servo loop. The extent of access determines gain. If the thumb is anaesthetic, access is barred. A movement can still be made, involving, as usual, activation of both α and γ cortical motoneurones, but it requires greater effort, of which the subject is aware. Although the spindles will contract, they cannot thereby give servo assistance, because the feedback pathway to the cortical α motoneurones is not open—at any rate in the experimental situation with which we are at present dealing, in which visual clues to the perturbations introduced are excluded, and there are no auditory clues. This scheme explains all our present results, but assigns no role to the spinal stretch reflex, although the existence of tendon jerks in healthy subjects, of a latency consistent only with a spinal reflex arc, proves that the spinal stretch reflex is competent to act in ordinary life.

Quite independently, however, of such interpretations in terms of specific receptor mechanisms, we are thus making five new experimental points: (1) that there *is* servo action, sensitive, brisk and so early as clearly to be automatic, in voluntary movements of the thumb, (2) that the gain of the servo loop is proportional to the force exerted, (3) that the gain of the servo loop is automatically compensated against contractile fatigue in the muscle, (4) that the latency of servo action is the same as the latency of the stretch reflex, and (5) that both servo action and the stretch reflex are heavily dependent on sensation in the thumb itself, as opposed to the muscle that moves it. The relevance of these findings to the interpretation of spastic and other pathological states of muscle in man will not escape notice.

Finally, while making acknowledgment to the Department of Trade and Industry, who, in part, financed the work with some such return in mind, it is appropriate to observe that whereas the servo theory of muscular contraction arose from the application of engineering ideas to physiology, it looks as if nature may be about to repay the compliment and reveal some novel mechanisms for control engineers to reflect upon.

Received April 20; revised June 7, 1972.

1 Merton, P. A., in *The Spinal Cord* (edit. by Malcolm, J. L., and Gray, J. A. B.), 247 (Churchill, London, 1953).
2 Eldred, E., Granit, R., and Merton, P. A., *J. Physiol.*, **122**, 498 (1953).
3 Matthews, P. B. C., *Physiol. Rev.*, **44**, 219 (1964).
4 Phillips, C. G., *Proc. Roy. Soc.*, B, **173**, 141 (1969).
5 Marsden, C. D., Merton, P. A., and Morton, H. B., *J. Physiol.*, **216**, 21 P (1971).
6 Hick, W. E., *Quart. J. Exp. Psychol.*, **4**, 11 (1952).
7 Hammond, P. H., in *Proc. Third Int. Conf. on Medical Electronics*, 190 (Institution of Electrical Engineers, London, 1960).
8 Merton, P. A., *J. Physiol.*, **114**, 183 (1951).
9 Granit, R., Holmgren, B., and Merton, P. A., *J. Physiol.*, **130**, 213 (1955).
10 Granit, R., *The Basis of Motor Control* (Academic Press, London and New York, 1970).
11 Vallbo, Å. B., *J. Physiol.*, **218**, 405 (1971).
12 Marsden, C. D., Merton, P. A., and Morton, H. B., *J. Physiol.*, **222**, 32 P (1972).
13 Merton, P. A., *Brit. Med. Bull.*, **12**, 219 (1956).
14 Phillips, C. G., Powell, T. P. S., and Wiesendanger, M., *J. Physiol.*, **217**, 419 (1971).
15 Carmichael, E. A., *Proc. Roy. Soc. Med.*, **40**, 553 (1947).
16 Dawson, G. D., *J. Neurol. Neurosurg. Psychiat.*, **10**, 141 (1947).
17 Halliday, A. M., *Brain*, **90**, 241 (1967).
18 Lance, J. W., and Adams, R. D., *Brain*, **86**, 111 (1963).
19 Kugelberg, E., and Widén, L., *EEG Clin. Neurophysiol.*, **6**, 503 (1954).