

1992
11 *The skeletomotor system*

JOHN T. CACIOPPO, LOUIS G. TASSINARY, AND
ALAN J. FRIDLUND

The principal function of the nervous system is the coordinated innervation of the musculature. Its fundamental anatomical plan and working principles are understandable only on these terms. (Sperry, 1952, p. 298)

Certain complex actions are of direct or indirect service under certain states of the mind, in order to relieve or gratify certain sensations, desires, etc.; and whenever the same state of mind is induced, however feebly, there is a tendency through the force of habit and association for the same movements to be performed, though they may not be of the least use. (Darwin, 1873/1872, p. 28, italics added)

There has been a marked hesitation in thinking of the motor system as having anything significant to contribute to higher mental processes.... The violinist Itzhak Perlman, in trying to play a difficult note raises his eyebrows (if it is a high note) and keeps them raised until the note has been played. His face and body perform a rich program of varied movements. Why, again? With few exceptions (Piaget, 1954), it is generally believed that these motions are secondary and ancillary. But suppose that a good part of musical memory is in fact lodged in these peculiar movements. Suppose that they are significant. (Zajonc & Markus, 1984, pp. 81 - 84)

11.1 INTRODUCTION

The skeletomotor system is the final common pathway through which humans interact with and modify their environment. The specificity and sophistication of the skeletomotor system enable the vast repertoires of adaptive reflexes and skilled behaviors associated with living organisms (cf. Smith and Kier, 1989). The electrophysiological signals emanating from the muscular system have been of interest for over four centuries due to the complexity of their organization and dynamics; their clinical applications; and their value as indices of and possible contributors to processes such as cognition, motivation, and emotion.

In this chapter, we provide an introduction to psychophysiological research on the skeletomotor system. We begin by reviewing the history of this research and by articulating some of the major issues, limitations, and advantages of surface electromyography (EMG). We then review briefly the physiological basis of the EMG, and we summarize and update the recent guidelines for surface EMG recording in humans by Fridlund and Cacioppo (1986). We continue with a discussion of the social context for EMG recording and of psychophysiological principles, paradigms, and applications that have emerged from research on the skeletomotor system. Due to the anatomy and electrophysiology of the skeletomotor system, many of the topics and issues addressed in this chapter are of general importance in psychophysiology.

Finally, a major theme in this chapter is that technical competence in recording EMG activity is necessary but not sufficient for securing scientifically meaningful data. This is because analyzing what underlies any muscular act is complex. Physiologically, similar limb displacements or feature distortions are often achieved by distinctly different muscular actions (e.g., Gans & Gorniak, 1980); control over these actions can be achieved peripherally through reflex arcs or centrally through the extrapyramidal or pyramidal mediation (e.g., Henneman, 1980a-c). At the behavioral level, muscular acts do not always occur as intended (e.g., as when one performs clumsily), are sometimes nonobvious (e.g., as when one "hides feelings" or inhibits an action), are often misleading as to goals (e.g., as when one deceives), and may not always be in the service of a single psychological endpoint (e.g., as when one lowers the brows in sadness vs. to communicate a point). For EMG signals to be of theoretical significance, therefore, one must consider the physiological, social, and inferential contexts in which these signals are acquired. Hence, in this chapter we address all three of these aspects of surface electromyography. Additional information on electromyography can be found in Basmajian and DeLuca (1985), Loeb and Gans (1986), Fridlund and Cacioppo (1986), Goldstein (1972), Strong (1970), and Lippold (1967).

11.2 HISTORICAL BACKGROUND

At least two distinct themes in the development of electromyography in psychophysiology can be identified. The first is the history of the physiology of the muscles, which derives from the writings of the early Greek philosophers and physicians, and of the field of neurophysiology, which can be traced to Francesco Redi's (1925/1671) deduction that the shock of the electric ray fish (the *Torpedo torpedo*) emanated from specialized muscle tissue (Wu, 1984). Redi's work provided early evidence that the muscles were a source of electricity (Basmajian & DeLuca, 1985). Within four decades, William Croone (1633-1684), through a bequest in his will, founded the revered Croonian lecture of the Royal Society on the physiology of "muscular motion." These annual lectures have been delivered each year for over 250 years; no other field of physiology has an older lectureship devoted to its advancement (Fulton, 1926).

The second theme relates to the theory guiding much of the psychophysiological research using EMG, which owes a debt to the work of such figures as Duchenne (1959/1867), Spencer (1855), Darwin (1873/1872), and James (1884, 1890), all of whom emphasized relatively subtle patterns of muscular actions as a way of characterizing and understanding organismic-environmental transactions more generally.

11.2.1 The history of muscle physiology

The history of muscle physiology can be broken into four partially overlapping phases (Fulton, 1926): (1) early history, fourth century B.C. to late 1700s; (2) electrophysiology, late 1700s to present; (3) thermodynamics and chemistry of contraction, mid-1800s to present; and (4) neural control of contraction, late 1800s to present.¹

11.2.1.1 Early history

The first period was characterized by both experimental and theoretical investigations into the causes of animal movement. In his books *De Motu Animalium* and *De*

The skeletomotor system

processu Animalium, Aristotle (384-322 B.C.) provided clear descriptions of coordinated motor acts (e.g., locomotion and the importance of the mechanism of flexion). Claudius Galen (131-201 A.D.), along with his many other discoveries in experimental physiology (see Cacioppo & Tassinary, chapter 1), may have been the first to observe that a muscle is still capable of contracting even after surgical removal from the body. Galen also was the first to note correctly that muscles can only contract or relax.

11.2.1.2 Electrophysiology

Francesco Redi (1925/1671) and his pupil Stephano Lorenzini (1678) were the first to dissect the torpedo ray and to conclude that the electric organ was a specialized muscle tissue (Wu, 1984). Giovanni Borelli (1680) proposed that the current emanating from this specialized muscle resulted from rapid contractions that produced numerous sharp blows and a numbing of anything touching it. It was not until 1762 that Pieter van Musschenbroek, the inventor of the Leyden jar (an early device used for storing electrostatic charge), noted that the pain incurred upon touching a charged Leyden jar and an electric eel were similar and suggested that the pain from the muscle contractions of the eel was also electrical. Contrary to the then prevalent theory that the torpedo's shock was mechanical in nature, Edward Bancroft (1769, as cited in Wu, 1984) and John Walsh (1773) conducted studies demonstrating that the eel and torpedo's shock could be transmitted through liquids. The notion that muscle contraction was electrical in nature continued to be doubted, however, because muscle contraction was not associated with the sparks that were commonly observed in studies of airborne electrostatic charge. Walsh, in 1776, was finally able to demonstrate that the current from an electric eel could produce a spark:

The strongest shocks of the gymnotus will pass a very short interruption in the circuit When the interruption is formed by the incision made by a pen-knife on a slip of tin-foil that is pasted on glass, and that slip is put into the circuit, the shock in passing through the interruption, will shew a small but vivid spark, plainly distinguishable in a dark room. (Cavallo, 1786, pp. 309-311)

Direct evidence for a relationship between muscle contraction and electricity was not obtained until the late eighteenth century when Luigi and Lucia Galvani (and, later, their nephew, Aldini) conducted a series of studies on muscular contractions evoked by the discharge of static electricity (Galvani, 1953/1791). The Galvanis also reported that the muscles of a frog's legs were depolarized merely by touching them with metals rods, and he noted that the intensity of the contraction depended on the type of metal used (Foley, 1954). Galvani interpreted these observations as meaning electricity was stored in the muscles. Alessandro Volta initially endorsed Galvani's hypothesis. However, upon replication and careful observations using a sensitive condensing electroscope he had invented earlier, Volta (1792/1816) argued that the muscle contractions did not result from current arising within the organism but rather from electrical current generated because dissimilar metals touched the muscle preparation. Although Volta was correct in asserting that electric current can arise from the contact of dissimilar metals, so too was Galvani's hypothesis that living cells produce electricity. Galvani (1952/1794) repeated his experiments, but to stimulate the muscles he used nerves from a severed spinal cord rather than a metal

arc. He again found that the muscles contracted when brought into contact with the severed nerve.² Alexander von Humboldt (1797) subsequently replicated Galvani's findings using a large variety of animals, but according to Fulton (1926, p. 38), the notoriety accorded Volta for his many inventions minimized the impact of these observations for almost 40 years.³

It was not until the early-nineteenth century that the galvanometer, a sensitive instrument for measuring electric currents, was invented. In 1833, Carlo Matteucci used a galvanometer to demonstrate an electrical potential between an excised frog's nerve and its damaged muscle. In 1841, the renowned physiologist Johannes Müller (1801–1858) handed Matteucci's recent publication to one of his students, and seven years later the student published the results of an extensive series of investigations on the electrical basis of muscular contraction (Du Bois-Reymond, 1849). Du Bois-Reymond provided the first evidence of electrical activity in human muscles during voluntary contraction. Du Bois-Reymond's classic experiment involved placing a blotting cloth on each of the hands or forearms and immersing them in separate vials of saline solution. Each of these "electrodes" was attached to a galvanometer. Du Bois-Reymond observed minute deflections of the galvanometer when the muscles in the arms and hands were flexed. He further reasoned that the impedance (nonconductive nature) of the skin made the small voltage changes emanating from the muscles especially difficult to detect. To reduce this impedance, he created a blister on each forearm, removed the blistered skin, and placed the electrodes over these raw regions. Du Bois-Reymond found that contracting the arms and hands now resulted in much larger deflections in the galvanometer.

11.2.1.3 Thermodynamics and chemistry

The study of the thermodynamics of muscle contraction owes a debt to another of Müller's students, Hermann Ludwig Ferdinand von Helmholtz (1821–1894). Fueled by the desire to abolish the notion of vital forces underlying muscular actions, von Helmholtz began an investigation into the chemical transformations occurring in frog muscle during contraction. Based on the recently proposed Law of Conservation of Energy (i.e., the First Law of Thermodynamics), von Helmholtz reasoned that the heat of combustion combined with the transformation of food material should produce a quantity of heat measurable at the muscle surface during contraction. By stimulating an isolated muscle through its nerve and employing a sensitive thermocouple, von Helmholtz was able to demonstrate a rise in temperature during contraction.

Von Helmholtz's demonstration not only provided the experimental basis for his classic paper on the conservation of energy, but also it proved instrumental in focusing subsequent investigations on the central problem of understanding the physiochemical processes involved in converting neural energy to mechanical work. A historical account of our understanding of the physical processes involved in muscle contraction can be found in Needham (1971).

11.2.1.4 Neural control

Based on experimental observations using electrical stimulation, muscle physiologists since the time of Galvani attributed graded muscular responses to graded variations in the intensity of the stimulation. Until late in the nineteenth century,

The skeletomotor system

many erroneously inferred from this high correlation between the intensity of exogenous electrical stimulation and the intensity of contraction that the actual size of the neural impulses was proportional to the stimulus intensity.

Experimental work at the turn of the century by Bowditch (1871), Gotch (1902), Lucas (1909), and others challenged this belief, strongly suggesting that graded muscular responses resulted from the firing of individual contractile units rather than from variation in the size of the nerve impulse. Direct evidence for the "all-or-none" character of the response of muscle fibers was not obtained, however, until the work of Pratt and his colleagues in the early 1900s (Pratt, 1917; Pratt & Eisenberger, 1919). They provided graded electrical stimulation to individual muscle fibers either individually or in small groups while simultaneously photographing the spatial displacement of mercury droplets sprinkled over the muscle surface. They directly observed through a microscope that additional fibers contracted coincident with each quantal step in the displacement of a mercury droplet.

11.2.2 Skeletomotor activation and patterning

Detecting myoelectric signals using surface electrodes remained difficult throughout the nineteenth and early twentieth centuries. Electrically stimulating a muscle cutaneously was considerably simpler, however, and gained wide attention. Perhaps best known for this work was Guillaume Duchenne, who used this technique to investigate the dynamics and function of intact skeletal muscles (Duchenne, 1959/1867). At this point, the two lines of history mentioned at the outset of this section dovetail, as Darwin (1873/1872) corresponded with Duchenne in an effort to test his observations about facial expressions and bodily gestures.

Darwin's (1873/1872) interest in muscular action grew from his belief that phylogenesis was continuous and that behaviors as well as biological structures were in part inherited. He focused on the expression of emotions in man and animals to illustrate the latter. Darwin suggested that expressive movements stemmed primarily from the inheritance of acquired habit, secondarily through what he termed antithesis, and thirdly through the adventitious wiring of the nervous system. Darwin's (1873/1872) mechanism of inheritance was primarily Lamarckian.

That some physical change is produced in the nerve-cells or nerves which are habitually used can hardly be doubted, for otherwise it is impossible to understand how the tendency to certain acquired movements is inherited. (p. 29)

This was not the only mechanism of origin of expressions and behaviors, however, as the processes of variation and natural selection were also embraced:

Nor must we overlook the part which variation and natural selection may have played; for the males which succeeded in making themselves appear the most terrible to their rivals, or to their other enemies, if not of overwhelming power, will on an average have left more offspring to inherit their characteristic qualities, whatever these may be and however first acquired, than have other males. (p. 104)

Although Darwin's (1873/1872) observations were limited to overt actions, he suggested that "whenever the same state of mind is induced, however feebly, there is a tendency through the force of habit and association for the same movements to be performed, though they may not be of the least use" (p. 28). This suggestion

presaged contemporary studies of the patterns of muscle contractions and facial actions that are undetectable to the naked eye (see Cacioppo, Martzke, Petty, & Tassinary, 1988).

The somatic components of William James's (1884, 1890) theory of emotions, James's (1890) ideomotor theory, and the various motor theories of thinking prevalent at the turn of the century (cf. McGuigan, 1978) further fueled interest in objective measures of subtle or fleeting muscle contractions. Among the more creative procedures used to magnify tiny muscular contractions was the placement of a flattened wine glass on the tongue to serve as a sensor during thought and of mechanical extensions from the wine glass to amplify the movements of the tongue (see McGuigan, 1979). However, accurate, reliable, and sensitive noninvasive recordings awaited the development of metal surface electrodes, vacuum tube amplifiers, and the cathode-ray oscilloscope early this century and the subsequent pioneering work of Edmund Jacobson (1925, 1927, 1930a-d, 1931a-c, 1932) on electrical measurements of muscle unit action potentials during imagery.

Jacobson's general approach has much to offer contemporary investigators. Briefly, he trained subjects in progressive relaxation, thereby obtaining very low basal levels of muscle tonus, and then ran subjects through a series of mental tasks. Clicks of a telegraph key were presented to mark the onset and offset of each task. Subjects were instructed to engage in a particular "mental activity" following the first click and to relax any muscular tensions present following the second. Tasks included "imagine throwing a ball," "imagine counting," "recall a poem," and "imagine the Eiffel Tower in Paris." Jacobson also employed a control procedure of instructing his subjects on some trials that "upon hearing the first signal do not bother to think." The results of these studies indicated that (1) EMG responses were evoked by these tasks, (2) these responses were minute and highly localized, and (3) these localized responses often occurred in the part of the body that one would use had the task called for an overt response (e.g., see Jacobson, 1932). This work was criticized primarily for not definitively achieving mentally quiescent comparison periods (e.g., Humphrey, 1951; Max, 1937), but successful replications of some of Jacobson's early work using different comparison tasks have been reported (McGuigan, 1978, chapters 6 and 7).

11.2.3 Issues in EMG research

Subsequent research using surface EMG has extended these early observations, documenting patterns of covert skeletomotor activity that differentiate both within and between emotional and cognitive processes (e.g., see reviews by Cacioppo & Petty, 1981a; Cacioppo, Losch, Tassinary, & Petty, 1986; Fridlund, 1988; Fridlund & Izard, 1983; McGuigan, 1978) as well as between normal and clinical populations (e.g., van Boxtel, Goudswaard, & Janssen, 1983; Malmo, 1975; Schwartz, Fair, Sall, Mandel, & Klerman, 1976; Whatmore & Ellis, 1959). The more important issues in Jacobson's early work and in subsequent psychophysiological research include the extent to which the EMG responses reflect: (1) specific or global activation; (2) phasic activation, tonic activation, or modulated thresholds for activation; and (3) characteristics of the stimulus (situation) or the individual (disposition). To address these issues, it has been necessary to monitor multiple measures of EMG activity. Furthermore, it has often been advantageous to monitor EMG responses across time using within-subjects designs; to use appropriate

The skeletomotor system

controls for assessing practice, sensitization, and carryover effects; to employ time-locked recording procedures; and to perform idiographic analyses at least when individual differences are of concern.

In addition, understanding the relative advantages and disadvantages of EMG in studies of psychology is important for its correct application and interpretation. Most of the striated muscles in the human body are bilaterally symmetrical in pairs (i.e., one on each side of the body), with estimates of the number of distinct muscles in the body ranging from 450 (Anson, 1966) to 639 (Khan, 1943) to 792 (Tomovic & Bellman, 1970). The general distribution of muscles across the body, as depicted in Anson (1966), is as follows: 37 bilaterally symmetrical muscle pairs in the head and face; 29 pairs in the neck and shoulder girth; 54 pairs in the shoulders, arms, and hands; 21 pairs in the spinal region; 15 pairs in the thoracoabdominal region; 9 pairs in the pelvic outlet; and 62 pairs in the hip, thighs, legs, and feet (Figure 11.1).

From a functional perspective, each striated muscle can be characterized as a linear actuator, with the potential states being limited to onset of contraction, offset of contraction, and relaxation (Tomovic & Bellman, 1970), wherein maximal muscle contraction results in a reduction in length to approximately 57 percent of that observed at rest (cf. Willis & Grossman, 1977). However, the structural arrangements of the striated muscles (e.g., as agonist-antagonist pairs, or through their interdigitation) expand dramatically the number of outcomes that can be achieved by this fairly simple system. The relatively small number of muscles in the head and neck, for instance, have been estimated to enable the encoding of some 6,000 to 7,000 appearance changes (Izard, 1971; see Figure 11.1 left panel). With this flexibility comes adaptability:

Speaking of muscles as linear actuators, it must always be kept in mind how inadequate the engineering performance criteria are when extended in any routine fashion to the biological world. For instance, we are so used to wheeled vehicles that we forget that there is a very close relation between their efficiency and road conditions. Animal locomotion, however, is practically independent of environmental constraints. This optimization criterion is therefore much more essential for the survival goal. (Tomovic & Bellman, 1970, p. 273)

This skeletomotor architecture and its flexibility of action are also the source of the major limitations in EMG studies of psychology or behavior. First, it is feasible to obtain measurements over only a small number of muscles in the human body in any given experiment. Yet because the action of the striated muscles is multiply determined, monitoring activity from a single site may only provide global or ambiguous information about the associated psychological or behavioral process. Ekman (1982) observed, for instance, that emotions, with the possible exception of happiness, cannot be identified by the activity of a single muscle: "Disgust might be measured by the activity of two muscles, and surprise by the activity of three. To measure anger, fear, or sadness, many muscles need to be measured" (Ekman, 1982, p. 79).

Second, many limb displacements and feature distortions can be achieved by the actions of different or differently activated striated muscles. Electromyographic responses may therefore appear unreliable if the focus is on the motor/behavioral output rather than the mechanisms by which these outcomes were achieved (Gans & Gorniak, 1980). Third, the imperfect selectivity of surface electrodes and the close proximity of the various striated muscles make it difficult to pinpoint exactly which muscles are contracting. Hence, when using surface electrodes, it is appropriate

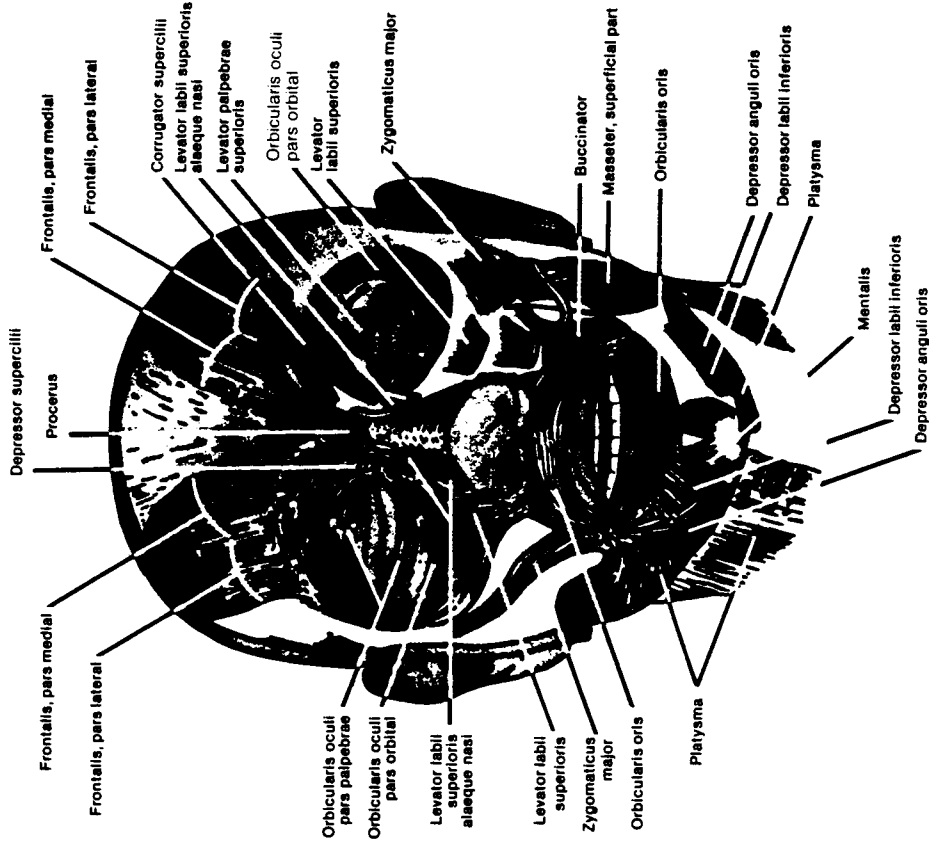
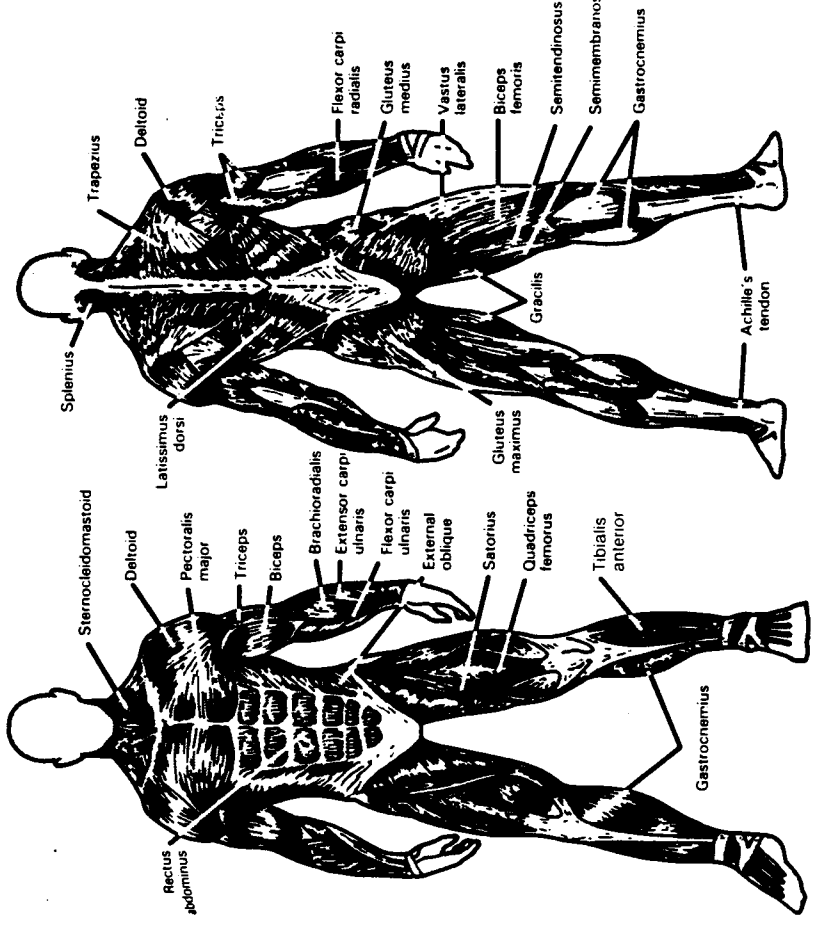


Figure 11.1. a: Schematic representation of selected facial muscles. Overt facial expressions of emotion are based on contractions of underlying musculature that are sufficiently intense to result in visibly perceptible dislocations of skin and landmarks. More common visible effects of strong contractions of depicted facial muscles include the following. Muscles of lower face: *depressor anguli oris*, pulls lip corners downward; *depressor labii inferioris*, depresses lower lip; *orbicularis oris*, tightens, compresses, protrudes, and/or inverts lips; *mentalis*, elevates chin boss and protrudes lower lip; *platysma*, wrinkles skin of neck and may draw down both lower lip and lip corners. Muscles of midface: *buccinator*, compresses and tightens cheek, forming a "dimple"; *levator labii superioris alaeque nasi*, raises center of upper lip and flares nostrils; *levator labii superioris*, raises upper lip and flares nostrils, exposing canine teeth; *masseter*, raises lower jaw; *zygomaticus major*, pulls lip corners up and back. Muscles of upper face: *corrugator supercillii*, draws brows together and downward, producing vertical furrows between brows; *depressor supercillii/procerus*, pulls medial part of brows downward and may wrinkle skin over bridge of nose; *frontalis, pars lateral*, raises outer brows, producing horizontal furrows in lateral regions of forehead; *frontalis, pars medial*, raises inner brows, producing horizontal furrows in medial region of forehead; *levator palpebrae superioris*, raises upper eyelid; *orbicularis oculi pars orbitalis*, tightens skin surrounding eye, causing "crow's-feet" wrinkles; *orbicularis oculi, pars palpebrae*, tightens skin surrounding eye, causing lower eyelid to rise. Descriptions are consistent with those in Daniels and Worthingham (1986). Ekman and Friesen (1978), Kendall and McCreary (1980), Izard (1971), and Weaver (1977). (From J. T. Cacioppo, J. S. Martzke, R. E. Petty, & L. G. Tassinari, 1988. Reproduced with permission.)



b: Schematic representation of major muscle groups of human body. Overt skeletomotor actions based on contractions of underlying musculature that are sufficiently intense to result in movements of skeletostucture. Front (ventral) view: *Sternocleidomastoid*, rotates face toward side opposite contracting muscle and, when paired muscles are contracted, flexes vertebral column to move neck and face toward chest; *deltoid*, abducts arm; *pectoralis major*, flexes, adducts, and medially rotates arm; *biceps brachii*, flexes and supinates arm; *triceps brachii*, extends forearm; *external oblique*, bends vertebral column laterally toward contracting muscle and, when paired muscles are contracted, compresses abdomen; *brachioradialis*, flexes forearm; *flexor carpi ulnaris*, flexes and adducts wrist; *extensor carpi ulnaris*, extends and abducts wrist; *sartorius*, flexes leg and flexes and rotates thigh laterally; *quadriceps femoris*, composite of four muscles (*rectus femoris*, *vastus lateralis*, *vastus medialis*, and *vastus intermedius*), which extend leg, and *rectus femoris*, which flexes thigh; *tibialis anterior*, dorsiflexes and inverts foot; *gastrocnemius*, plantar flexes foot; *rectus abdominis*, flexes vertebral column. Back (dorsal) view: *splenius*, rotates head toward side same size as contracting muscle and, when paired muscles are contracted, extends head; *trapezius*, adducts, rotates, and elevates scapula (back part of shoulder) and extends head; *latissimus dorsi*, extends, adducts, and rotates arm medially; *gluteus medius*, abducts and rotates thigh medially; *gluteus maximus*, extends and rotates thigh laterally; *deltoid*, abducts arm; *flexor carpi radialis*, flexes and abducts wrist; *hamstrings*, composite of three muscles (*biceps femoris*, *semitendinosus*, and *semimembranosus*), which flex leg and extend thigh; *gracilis*, flexes leg and adducts thigh; *vastus lateralis*, extends leg. Descriptions are consistent with those in Daniels and Worthingham (1986), Langley, Telford, and Christensen (1974), Schmidt and Thews (1983), and Tortora (1983). Note on terminology: *Abductor* moves bone away from midline of body; *adductor* moves bone closer to midline; *levator* produces upward movement; *depressor* produces downward movement; *sphincter* decreases size of opening; *flexor* usually

(contd. on next page)

only to refer to EMG signals as reflecting activity from sites or muscle regions (e.g., "*corrugator supercilii* muscle region"). Fourth, surface EMG recording, although noninvasive, can be obtrusive and potentially reactive. Electrodes attached to the surface of the skin and leads traveling to preamplifiers, for instance, can restrict an individual's movement or make the individual tense or self-conscious. Finally, until recently there was no theoretically derived or empirically verified standard for the placement of surface electrodes to detect activity in a particular region of the face or, to some extent, the body. The absence of such a standard weakened comparisons across laboratories or across individuals and sessions within laboratories.

Progress has been made in overcoming many of these limitations (e.g., Fridlund & Cacioppo, 1986; Tassinari, Cacioppo, & Geen, 1989; Tassinari, Cacioppo, Geen, & Vanman, 1987), and this progress is reviewed in the sections that follows. In addition, surface EMG recording offers several unique advantages that complement the study of overt behavior through traditional means (see, also, Cacioppo & Petty, 1983; Fridlund, 1987). First, EMG responses, in contrast to measures such as response latencies or verbal reports, can be collected continuously without the individual's attention or labor. Second, the detection and quantification of EMG signals over a muscle region can be performed with the assistance of computers more sensitively, reliably, and quickly than can fine-grain analyses of overt behavior. Third, analyses of subtle somatic patterns and their time course may provide a means of differentiating underlying mechanisms of control over overt behaviors that are visibly identical.

Fourth, many subtle psychological (e.g., emotional) processes or events are not accompanied by visually perceptible actions or significant visceral changes, and these facts have hindered theory and research of psychological processes (e.g., Graham, 1980; Rajecki, 1983; cf. Cacioppo, Petty, et al., 1986). Darwin (1873/1872) recognized this limitation in the study of emotional expressions, stating that "the study of expression is difficult, owing to the movements being often extremely slight, and of a fleeting nature" (p. 12). It is now clear, however, that fast or low-level changes in EMG activity can occur without leading to any visible limb displacement or feature distortions on the surface of the skin. Facial expressions, for instance, result from movements of facial skin and connective tissue due to the contraction of facial muscles that create folds, lines, and wrinkles in the skin and the movement of facial landmarks such as the brows and corners of the mouth (e.g., Ekman & Friesen, 1978; Izard, 1971; Kinn, 1984). Although muscle activation must occur if these facial distortions are to be achieved (see Figure 1.1), it is possible for muscle activation to occur in the absence of any overt facial action if the activation is weak or transient or if the overt response is aborted. This holds for nonfacial striated muscles as well (e.g., see Coles, Gratton, Bashore, Eriksen, & Donchin, 1985).

In the face, this is due to the structure and elasticity of the facial skin, facial sheath, adipose tissue, and facial muscles. The muscles of expression are attached to

decreases anterior (or, in a few cases, posterior) angle at joint; *extensor* usually increases anterior (or, in a few cases, posterior) angle at joint; *supinator* turns palm upward or anteriorly; *pronator* turns palm downward or posteriorly; *dorsiflexor* flexes foot at ankle joint; *plantar flexor* extends foot at ankle joint.

The skeletomotor system

other muscles, bones, or a facial sheath below the surface of the facial skin and adipose tissue; not unlike a loose chain, the facial muscles can be pulled a small distance (i.e., contracted slightly) before exerting a significant force on the object to which they are anchored (cf. Tassinari et al., 1989). In addition, the elasticity of the facial sheath, facial skin, and adipose tissue forms a low-pass mechanical filter, attenuating the visible effects of very rapid contractions (Fridlund, 1987; cf. Fridlund & Cacioppo, 1986).

In summary, measures of EMG and of observable muscular actions each have unique advantages and disadvantages. Neither is necessarily better or more capable of capturing completely the information provided by the other. A general congruence between the results based on EMG recordings and those obtained through fine-grain analyses of overt behavior is to be expected given the physiological basis of the surface EMG (see what follows). Therefore, the wealth of information that exists regarding nonverbal behavior and displays during such processes as thinking, communication, deception, and emotion (e.g., see Scherer & Ekman, 1982) provides a rich theoretical resource for research on subtler, more fleeting responses or on underlying mechanisms.

Yet EMG recordings and fine-grain behavioral observations do not coincide completely. As noted, EMG recordings can reveal muscular activity or patterns of activity too small or fleeting to evoke detectable movements or whose corresponding muscle contractions are counteracted by contraction of an antagonist. Ekman (1982) reported an interesting, illustrative study conducted by Ekman, Schwartz, and Friesen. Surface EMG recordings and high-quality videorecordings were secured simultaneously as individuals deliberately intensified the contraction of specific facial muscles (the *corrugator supercilii* and *medial frontalis*). Results revealed that measurements of facial feature distortions using Ekman and Friesen's (1978) Facial Action Coding System (FACS) and measurements from surface EMG over these muscle regions were highly correlated ($r = +.85$). Nevertheless, reliable EMG signals emerged at levels lower than could reliably be detected visually. These results speak well both for the validity of facial EMG measurement and for the possibility of tracking at least limited features of moment-by-moment psychological processes even in the absence of a visually detectable motor response (see, also, Cacioppo & Petty, 1979a; Cacioppo et al., 1988; Fridlund, Schwartz, & Fowler, 1984; Schwartz, 1975). Moreover, because individuals may be less likely to control momentary or minute muscular contractions that do not result in observable actions, discrepancies between EMG responses underlying covert versus overt responses may be of special interest (cf. Cacioppo, Petty, et al., 1986). We return to these issues following a discussion of the technical aspects of the surface EMG.

11.3 PHYSICAL CONTEXT

Almost three quarters of a century ago, Baines (1918, cited in Basmajian & DeLuca, 1985) argued that appropriate technical considerations should precede the collection or interpretation of data related to electrophysiological phenomena. In their survey of the EMG literature, Basmajian and DeLuca (1985) strongly concurred, observing that "this call still echoes among the numerous abuses that have been promulgated throughout the past seven decades" (p. 6). Although fascination or concern with technical issues and "state-of-the-art"

equipment can be overdone, an understanding of the physiological system one is studying and the bioelectrical principles underlying its responses serve several important purposes. These include the (1) intimation or stimulation of theory and development of operational definitions and procedures; (2) discrimination of signal from artifact; (3) safety of the individuals involved; (4) digital data acquisition and analysis and derivation of descriptive parameters that are reliable and valid representations of the physiological events of interest; and (5) guidance of inferences based on physiological data (see, also, Cacioppo & Tassinari, chapter 1; Cacioppo & Tassinari, 1989). In this section, therefore, we review the physiological basis of the surface EMG, and we outline principles and technical issues involved in obtaining valid measures of EMG activity.

11.3.1 Anatomical and physiological basis of the surface electromyogram

The striated muscles in the human body are capable of different types of muscular activity. They have tone, maintain postures, make reflexive movements in response to sensory stimuli, actuate spontaneous rhythmic contractions, and produce both spontaneous and voluntary contractions that can occur independent of immediate external stimuli (Willis & Grossman, 1977). A schematic of the central organization and control of the several hundred skeletomuscles in the human body is presented in Figure 11.2. A detailed description of the central organization and control of the motor system, although important, is beyond the scope of the present chapter. Interested readers can consult Henneman (1980a-c), Kandel and Schwartz (1985), and Willis and Grossman (1977).

At the peripheral level, one finds that each striated muscle is innervated by a single motor nerve whose cell bodies are primarily located in the ventral horn of the spinal cord or, in the case of the muscles of the head, in the cranial nerves of the brain stem. All behavior—that is, all actions of the striated muscles, regardless of the source of the neural signal (e.g., reflex arcs, pyramidal neurons emanating from the motor cortex)—result from neural signals traveling along these motor nerves. For this reason, the set of lower motor nerves has been designated the final common pathway (Sherrington, 1923/1906).

Closer inspection of these neuromuscular units reveals that the typical striated muscle consists of hundreds or thousands of separate, elongated muscle fibers bound together by a sheet of connective tissue. The motor nerve traveling to the muscle consists of numerous individual motoneurons, which as a collective are referred to as a motoneuron pool. Each motoneuron axon divides into a number of small branches, termed *axon fibrils*, just before reaching the muscle; each axon fibril, in turn, forms a junction, called a motor end plate, on an individual muscle fiber (Figure 11.3). Each motoneuron innervates a number of interspersed muscle fibers within a muscle, and each muscle fiber is usually innervated by only one motoneuron. An important functional consequence of this structure is that muscle fibers do not contract individually but rather there is a concerted action by each set of muscle fibers innervated by a single motoneuron. Therefore, the most elementary functional unit within the "final common pathway" is the motoneuron cell body, its axon, its axon fibrils, and the individual muscle fibers innervated by these axon fibrils. This functional, physiological entity is called a motor unit (see Figure 11.3).

The axons of the motoneurons within a motoneuron pool vary in diameter, and this structural feature also has important functional consequences. Generally, the

The skeletomotor system

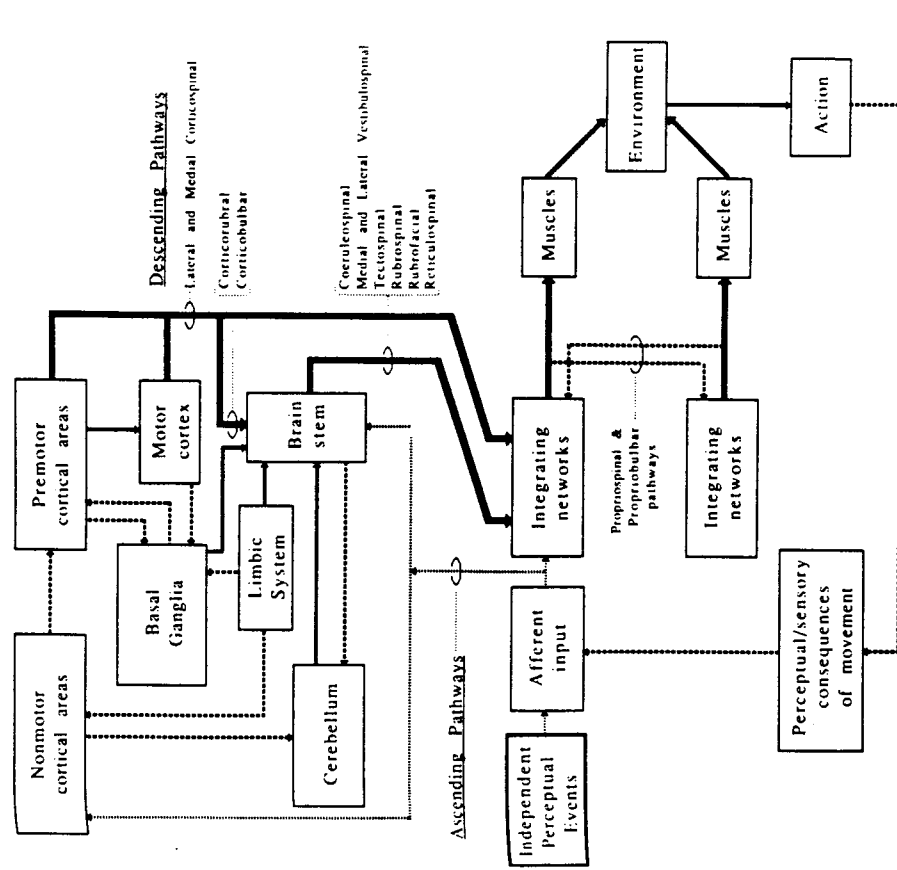


Figure 11.2. Major components of motor system. Solid lines represent efferent pathways, heavy dashed lines represent feedback pathways, and light dashed lines represent afferent pathways. Converging arrows do not necessarily imply convergence on same individual neurons. Crossing of pathways as well as details of specific connections within brain areas are not indicated. Note that arrows denote strong influences; they do not imply direct (monosynaptic) connections. Note also that list of descending pathways is representative; it is not meant to be exhaustive. The thalamus and hypothalamus have been omitted for clarity. (Modified and redrawn from Figure 33.3 of Ghez, 1985, based on Hinsey, 1940, and Kuypers, 1982.)

smaller the diameter of a motoneuron, the smaller the number of axon fibrils and consequently the smaller the number of muscle fibers it innervates. Hence, activation of muscle via small motoneurons produces smaller and more precise actions than activation of the same muscle by the depolarization of large motoneurons. In addition, the smaller the diameter of the motoneuron, the lower tends to be the critical firing threshold of its cell body and the more fatigue resistant (i.e., the greater the glycolytic capacity) the muscle fibers innervated by the

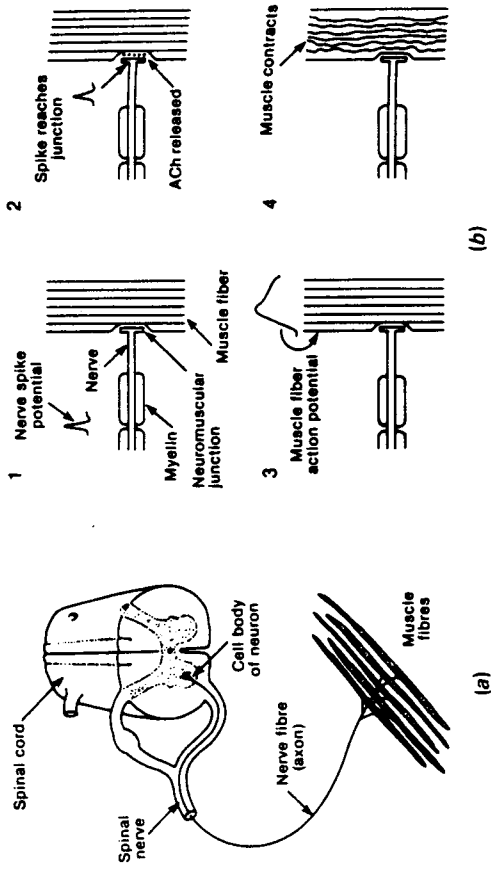


Figure 11.3. *Left panel (a)*: Schematic of motor unit. (From Basmajian & DeLuca, 1985. Reproduced with permission.) *Right panel (b)*: Sequence of events involved in transmission across neuromuscular junction: 1, nerve action (spike) potential approaches junction; 2, spike reaches junction and triggers release of ACh; 3, ACh acts on muscle fiber to produce muscle action potential; 4, muscle contracts. (From Thompson, 1967. Reproduced with permission.)

motoneuron. These relationships constitute the size principle (Henneman, 1980b), and it contributes to the graded contraction and coordinated actions of muscles. More specifically, the force of contraction produced by a muscle is attributable to small motoneurons initially discharging intermittently and then discharging more frequently. Stronger muscle contractions are attributable to the depolarization of increasingly large motoneurons within the motoneuron pool. As muscle contraction approaches maximal levels, further increases in contraction are again attributable to the individual motoneurons firing more quickly (Petrofsky & Phillips, 1982). If maximal contraction is required for an extended period, the physiochemical mechanisms underlying muscle contraction are unable to sustain brief refractory periods, and muscular quivering (tetany) develops.

The ratio of muscle fibers to motoneurons (innervation ratio) varies even more dramatically across than within muscles and is related to the precision of the movement of the particular muscles involved. Consistent with the principles outlined in the preceding, muscles with low innervation ratios (i.e., few muscle fibers per motoneuron and, hence, small motor units) are capable of producing actions more rapidly and with greater precision than are muscles with high innervation ratios. For example, the extrinsic muscles of the eye, which are capable of very fast and fine movements, have innervation ratios around 10:1, whereas the more slowly and grossly acting postural muscles have innervation ratios of around 3,000:1 (Basmajian & DeLuca, 1985).

The depolarization of a motoneuron results in the quantal release of acetylcholine at motor end plates (see Figure 11.3). The activating neurotransmitter acetylcholine is quickly metabolized by the enzyme cholinesterase so that continued efferent

discharges are required for continued propagation of muscle action potentials (MAPs) and fiber contraction. Nonetheless, the transient excitatory potential within a motor end plate can lead to a brief (e.g., 1-ms) depolarization of the resting membrane potential of the muscle cell and an MAP that is propagated bidirectionally across the muscle fiber with constant velocity and undiminished amplitude. In the process, the physiochemical mechanism responsible for muscle contraction is activated (cf. Loeb & Gans, 1986). As the MAP travels along the muscle fiber, a small portion of this electrical activity passes through the extracellular fluids to the skin. It is these voltage changes that constitute the major portion of the surface EMG signal. Thus, the EMG does not provide a direct measure of tension or muscular contraction (mechanical events) but rather the electrical activity that accompanies these mechanical events.

Of course, the voltage changes that are detected in surface EMG recording do not emanate from a single MAP but from MAPs traveling across many muscle fibers within a motor unit (i.e., motor unit action potential, or MUAP) and even from MAPs traveling across several to hundreds of thousands of motor fibers due to the activation of multiple motor units. The surface EMG represents the aggregated electrical signals that reach the skin at a given moment in time. Not only are the details of the individual MAPs lost, but so too are the precise muscular origins. Reliable and sensitive information about the aggregate actions (or inactions) of motoneuron pools across time can nonetheless be obtained by careful attention to the elements of EMG recording and analysis (e.g., see Cacioppo, Marshall-Goodell, & Dorfman, 1983; Lippold, 1967; Petrofsky & Phillips, 1982). It is to these topics that we turn next.

11.3.2 Physical recording basis

The physical components for EMG recording are depicted schematically in Figure 11.4. These components include detection of the bioelectrical signals emanating from the subject; preamplification and preliminary signal conditioning; signal amplification, display, digitization, and storage; EMG signal representation; and inferential statistical analysis. In this section, we survey common problems of and standard methods for dealing with each of these physical components of EMG recording.

11.3.2.1 Detecting the bioelectrical signals

As outlined in the preceding section, the EMG signal emanating from the muscle is a quasi-random train of motor unit action potentials discharged by the contraction of striate muscle tissue. The signal train is characterized by a frequency range of several hertz to over 1 kHz and by amplitudes ranging at the surface of the skin from fractions of a microvolt to a few thousand microvolts. These frequency and amplitude characteristics are broader than most bioelectrical events of interest to psychophysicists, and they overlap both a variety of nonmuscular physiological responses (e.g., electroencephalographic, cardiodynamic) and the ubiquitous external AC signals that power most laboratory lights, electrical transformers, and equipment (see Marshall-Goodell, Tassinari, & Cacioppo, chapter 4). Consequently, detection of high-quality EMG signals from a localized muscle region requires careful attention to noise reduction and grounding practices (to eliminate extraneous

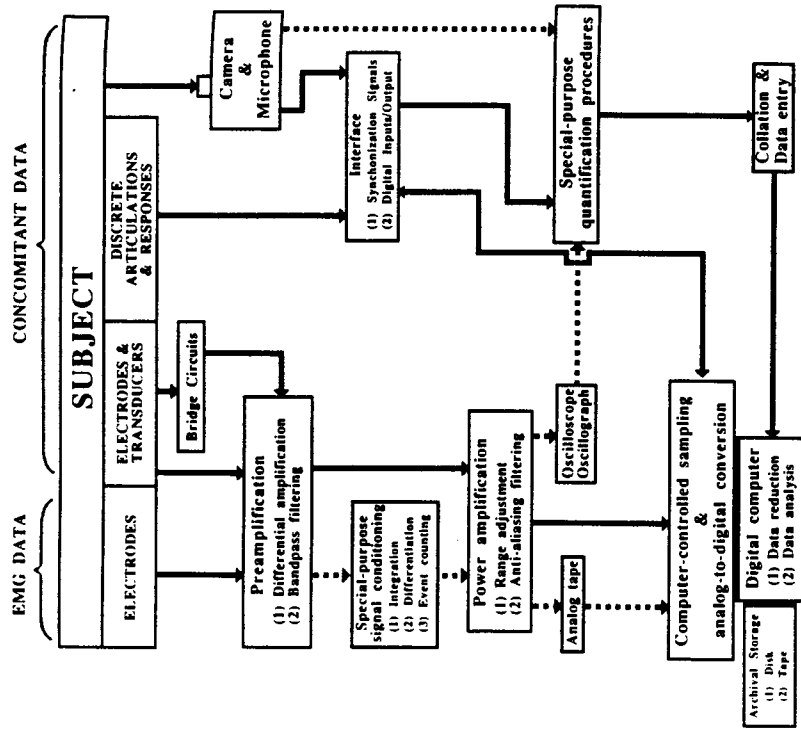


Figure 11.4. Major components of laboratory system for detection of EMG signals; preamplification and preliminary signal conditioning; power amplification, digitization, display, and data storage; and data processing (signal representation, statistical analyses). (Adapted from Cacioppo, Marshall-Goodell, & Gormezano, 1983, Loeb & Gans, 1986, and Shapiro & Crider, 1969.)

electrical noise), electrode site preparation and placement (to minimize the detection of irrelevant bioelectrical signals), and appropriate differential preamplification and preliminary signal conditioning (to further enhance signal-to-noise ratio).

Noise reduction and grounding procedures. Loeb and Gans (1986) define noise as any unwanted signal. The specific bioelectrical signal of interest can be obscured by noise from: (1) external electrical sources (e.g., through capacitive or inductive coupling from AC lines, movement of electrode interface or leads, electrical cross-talk in the recording componentry); (2) physiological responses whose frequency/amplitude characteristics overlap those for EMG recording (e.g., EKG, EEG); (3) EMG signals emanating from muscles whose actions are not of immediate interest at the target-recording site (i.e., cross-talk from somatic actions within the organism); and (4) EMG signals from target sites that result from "irrelevant" actions

one tried to minimize or eliminate by the design of the experiment (e.g., movement artifacts). It is little wonder that noise is the "bane of the electrophysiologist's existence" (Loeb & Gans, 1986, p. 21).

The most problematic extraneous electrical noise in the laboratory is narrow-band noise because it arises from several common sources, it radiates through walls and air, and its frequency range overlaps that of the EMG signal. Sixty-hertz noise emanates from AC power lines, lights, relays, and transformers. Although many EMG preamplifiers include a notch filter to attenuate 60-Hz noise, the filter is neither completely effective nor selective. That is, notch filters attenuate frequencies to a varying degree on both sides of 60 Hz—a bandwidth that can represent a significant portion of the total power in an EMG signal. It is therefore preferable to minimize electrical noise prior to amplification. This can be done through appropriate placement and shielding of equipment and careful grounding of the subject and equipment (Bramisley, Bruun, Buchthal, Guld, & Petersen, 1967; Strong, 1970).

Televisions, video monitors, and computer terminals all use cathode-ray tubes (CRTs), and they generate high-frequency electrical noise (ranging from 15 kHz to several hundred kilohertz) from the transformers used for CRT beam deflection. These electrical signals can be eliminated by placing the devices (and any unshielded AC power cords) at least 2 ft away from EMG electrodes, electrode leads, data lines, and equipment. In addition, the use of coaxial cable with shielded connectors is recommended to minimize the leakage of radio-frequency (RF) noise. Fridlund and Cacioppo (1986) and McGuigan (1979) describe procedures for measuring laboratory electrical noise sources.

Wide-band, or white-noise, is usually attributable to Brownian motion (i.e., atoms and molecules randomly colliding with each other) in electronic devices. This noise is minimized by keeping electrode impedances low and amplifier filters set tightly to the proper bandwidth (Loeb & Gans, 1986; Strong, 1970).

Biological noise includes all endogenous signals that are not part of the bioelectrical signal of interest. The placement of a ground electrode does not prevent artifacts from cross-talk. This is because the impedance into the ground electrode tends to exceed the access impedance of the volume-conductive tissues between the spurious current source and the recording electrode (Loeb & Gans, 1986; see Marshall-Goodell et al., chapter 4). In the sections that follow, we outline several procedures for minimizing biological noise. Additional details and discussions are provided by Loeb and Gans (1986), Strong (1970), and Basmajian and DeLuca (1985).

The subject affixed with electrodes should also be grounded at one and only one point on her or his body. If multichannel recordings are being made, then the ground planes for the recording channels should be strapped together rather than using separate grounds for each. Similarly, all equipment in the laboratory should be grounded at exactly one point, and any equipment touched by the subject should be nonconductive. These procedures further minimize 60-Hz noise in the recordings and enhance the safety of the subject by eliminating the possibility of ground loops (i.e., unintended current flow due to imperfect grounding).

Electrode selection and placement. Psychophysicists typically use surface rather than needle electrodes for EMG recording. This is due to the noninvasive nature of surface recording and to the research questions asked (thus far by

psychophysicologists (i.e., the usual interest is in muscles rather than motor units within muscles). Surface EMG electrodes are less sensitive to exact anatomical placement since they detect the MAFs from a cluster of motor units rather than a single unit. As previously noted, the discrete electrical discharges from individual MAFs summate spatially and temporally during motor unit recruitment to yield an aggregate that reflects the action of motoneuron pools. This aggregate response develops in an orderly manner from the individual MAFs - at least in a probabilistic sense - such that, generally, progressively larger motoneurons are added to or progressively smaller units subtracted from the total output from a motoneuron pool (Henneman, 1980a). Consequently, surface EMG recordings correlate well with the overall level of contraction of muscle groups underlying and near the electrodes, especially when limb movement is constrained and contractions are neither minimal nor maximal (Lawrence & DeLuca, 1983; Lippold, 1967).

Because most EMG amplifiers are AC coupled (see what follows), the electrical stability of the electrodes is not as important as, for instance, when recording skin conductance (e.g., see Dawson, Schell, & Fillion, chapter 10). Nonpolarizing electrodes such as silver-silver chloride (Ag/AgCl) electrodes are nevertheless preferable to stainless steel or other alloys due to variations in amplifier design, minimization of electrode artifact, and special problems inherent in detecting low-level signals (Fridlund, Tassinary, & Cacioppo, 1988).

Surface electrodes are available in a variety of sizes. Electrodes with small detection surfaces and housings allow closer interelectrode spacing and consequently higher selectivity.⁴ Factors such as the electrode size, electrode positioning, and interelectrode distance over a particular site can affect the detected EMG signals and, hence, should be held constant across experimental conditions. Fridlund and Cacioppo (1986) found that electrodes with 0.5 cm diameter Ag/AgCl detection surfaces and 1.5 cm diameter housings are used commonly for limb and trunk EMG sites, and miniature electrodes with 0.25 cm diameter Ag/AgCl detection surfaces and 0.5 or 1.0 cm diameter housings are used when greater recording selectivity is required, such as for facial EMG sites. Ad hoc recommendations were also offered for the use of a 1.0 cm interelectrode spacing for 0.25 cm electrodes and 1.5 cm spacing for the 0.5 cm electrode in bipolar EMG recording unless an explicit rationale exists for employing a different interelectrode spacing. Only closely spaced electrodes and differential amplification can yield spatially selective EMG recordings. Using widely spaced electrodes or electrodes that are not aligned with respect to the underlying striate muscle makes sense only if one is not interested in spatially selective EMG recording.

Specification of surface electrode placements over target muscle groups is important to ensure that findings are comparable across individuals, sessions, or laboratories. Several studies offer empirically and anatomically derived recommendations for EMG recording for facial, masticatory, and articulatory muscle activity using subdermal electrodes (e.g., Compton, 1973; Fridlund & Cacioppo, 1986; Isley & Basmajian, 1973; O'Dwyer, Quinn, Guitar, Andrews, & Neilson, 1981; Seiler, 1973; Vitti et al., 1975), and additional studies have examined the reliability of EMG measurements in relatively large, well-defined muscles (Gans & Gorniak, 1980; Komi & Buskirk, 1970). This research supports two principles of electrode orientation for differential recording over a given muscle region: (1) electrodes should be arranged to span maximally the gradient desired (e.g., in line with the underlying muscle whose activity is of interest) to maximize the recording of its activity and

The skeletomotor system

(2) electrodes should be arranged distal and/or perpendicular to gradients of extraneous signal sources (e.g., proximal muscles) to attenuate the recording of their activity. At the present time, only a handful of studies have established optimal electrode sites for bipolar surface EMG recording. Factors that must be balanced in implementing these principles in specifying electrode sites include (1) proximity of a proposed site to underlying muscle mass with minimal intervening tissue or interfering signals (such as from the electrocardiogram); (2) position of electrodes relative to muscle tissue fiber size, location, and orientation (e.g., electrodes generally should be aligned parallel to the course of the muscle fibers to maximize sensitivity and selectivity to the muscle of interest); (3) avoidance of straddling the motor end plate region; (4) ease of location of sites via anatomical landmarks that show relative uniformity across individuals; (5) ease of electrode attachments to these sites without undue obstruction of vision or movement or problems from skin folds, bony obstructions, and so on; and (6) minimizing cross-talk from proximal deep and superficial muscles (Fridlund & Cacioppo, 1986; Tassinary et al., 1989).

Davis (1952) was the first to offer recommendations for recording the muscle activity involved in movements of the limbs, neck, jaw, lower lip, and eyebrows. Davis, however, presented no evidence for the choice of particular sites over others and on occasion based placements on incorrect anatomical assumptions (cf. Fridlund & Cacioppo, 1986); consequently, only those regarding limb and neck placements are currently recommended (cf. Zipp, 1982).

Tassinary et al. (1989) provided relevant data for the *corrugator supercilii*, *depressor supercilii*, and *zygomaticus major* muscle regions, regions that have proven informative in studies of emotion. Based on anatomical data regarding the location of these muscles (see Figure 11.1a), several experiments were conducted to isolate the sites for surface EMG recording that met the six criteria outlined in the preceding. Subjects posed a series of facial actions and expressions twice while facial EMG activity was sampled. The activity of a specific muscle or set of muscles was verified with visible coding using the FACS (Ekman & Friesen, 1978). The surface recording sites identified as providing both sensitive and relatively selective measures of activation of specific muscle regions are illustrated in Figure 11.5.

Tassinary et al. (1987) provided relevant data for recording over the perioral muscle region and, in particular, for detecting silent language processing. Five sites in the perioral region were compared for their ability to differentiate facial actions due to the activation of discrete facial muscles in the perioral region. In addition to four electrode sites placed over the *mentalis*, *orbicularis oris superior*, *orbicularis oris inferior*, and *depressor anguli inferioris* (see Figure 11.1), a standard electrode placement (i.e., "chin") recommended by Davis (1952) to measure general perioral EMG activity was included. Activation of specific muscles or sets of muscles was again achieved by poses and was verified using the FACS. Following their correct performance of the poses on two separate occasions, subjects were asked to read affectively neutral passages silently, subvocally, and aloud. Only EMG responses recorded over the *mentalis*, *orbicularis oris inferior*, and chin muscle regions differentiated quiescent baseline activity from silent reading. Of these three sites, only the site depicted in Figure 11.5 for the *orbicularis oris inferior* muscle region demonstrated high discriminant validity when poses activating this versus proximal superficial muscles were contrasted.

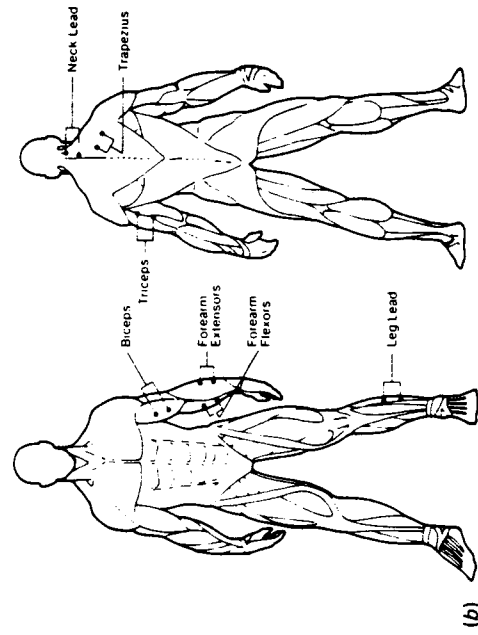
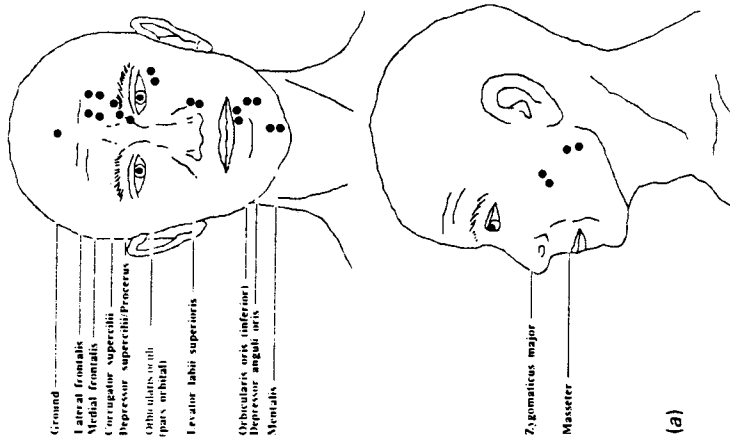


Figure 11.5. Atlas of EMG electrode placements for surface differential recording over selected facial muscle regions (a) and limb, body, and neck muscle regions (b). Detailed descriptions of electrode placements relative to physical landmarks are provided by Tassinari, Cacioppo, and Geen (1989), Fridlund and Cacioppo (1986), and Zipp (1982).

The skeletomotor system

Site preparation. Surface EMG electrodes can be attached to the skin using double-stick adhesive collars. A conductive medium (paste or gel) is used between the electrode surface and skin. This medium serves several functions: (1) stabilizing the interface between the skin and each electrode surface to minimize movement artifacts or alterations in recording area; (2) reducing interelectrode impedances by forming a conductive pathway across the hornified layers of the skin; and (3) stabilizing the hydration and conductivity of the underlying skin.

Prior to the application of the conductive medium and electrodes, the electrode site on the surface of the skin is usually cleaned to remove dirt and oil, and the site is typically abraded to lower interelectrode impedances to 5 or 10 k Ω . The electrodes are then commonly affixed in a bipolar configuration, as illustrated in figure 11.5. The placement of the ground electrode close but not directly adjacent to the EMG sites being monitored can also help minimize extraneous electrical noise in EMG recording. As already noted, however, only one ground electrode should be used, and all inputs should be configured to share this ground. Finally, to avoid obstructing movement due to the attachment of surface electrodes, thought should be given to the orientation of electrode collars and wires. Electrode wires, for instance, can be draped and secured using surgical adhesive tape to minimize distraction, annoyance, or obstruction of movement or vision.

11.3.2.2 Preamplification and preliminary signal conditioning

Electromyographic signals are "small" in two ways: They have low voltage and low current. As Loeb and Gans (1986) note, low-current signals coming from relatively high source impedances such as electrodes are particularly sensitive to noise picked up from external, extraneous signals (i.e., narrow-band noise), irrelevant, internally generated signals (i.e., biological noise) and the random motion of atoms and molecules within the amplification circuitry (i.e., wide-band noise). An amplifier supplies both voltage gain (turning low into high voltages), which can be controlled by the investigator, and current gain, a function of the ratio of the input and output impedances of the amplifier. The current gain of EMG preamplification is intentionally approximately zero to minimize source drain at the electrode site.

Electromyographic signals are amplified using differential amplifiers wherein the difference signal between two electrodes (with respect to a third, ground electrode) is amplified and carried through the signal-processing chain (Faulkenberry, 1977). Any bioelectrical or extraneous electrical signal that is common to both electrodes (the "common-mode" signal) is therefore attenuated (see Marshall-Goodell et al., chapter 4).

The most commonly used method of recording EMG signals is bipolar, in which electrode pairs are aligned parallel to the course of the muscle fibers. This alignment, coupled with the *common-mode rejection* of differential amplification, produces relatively sensitive and selective recording of the activity of the underlying muscle groups (Basmajian & DeLuca, 1985; see also Cooper, Osselton, & Shaw, 1980, chapter 3). Functional descriptions and circuit diagrams for differential EMG recording apparatus can be found in Fridlund and Fowler (1978) and Fridlund, Price, and Fowler (1982).

The second method of recording is monopolar, single ended, or common reference, which involves the placement of one EMG electrode over each target site (i.e., muscle group) of interest. The difference signal between the activity at each

target site and a common reference electrode (which, in theory, is in contact with an isoelectric site on the subject's body) is differentially amplified and carried through the signal-processing chain. Common reference recording is characterized by (1) a much more general pickup region than bipolar recording and (2) sensitivity to variations in the absolute level of electrical activity (assuming the ground electrode reflects an isoelectric state). Bipolar recording, in contrast, is sensitive to variations in the gradient of electrical activity between the two active electrodes. Due to these distinctions, the selection of the common reference or the bipolar method depends entirely on the question posed by the investigator (e.g., see Tassinari et al., 1989; Marshall-Goodell et al., chapter 4).

Regardless of the approach adopted, the signals recorded across time can be conceptualized as constituting a waveform. Any waveform can be represented in terms of the linear addition of pure sinusoidal waveforms of various frequencies, and this is the basis of Fourier analysis. A spectral analysis of a time series of raw EMG amplitudes, for instance, describes the time series by the weighting factors needed to synthesize it from a series of sinusoids that are harmonically related (see Porges & Bohrer, chapter 21). This is an important analytic procedure in surface EMG recording because by identifying the range of frequencies containing most of the energy of the signal, an investigator can better know the amplifiers, electrodes, and passband to employ.

A schematic representation of a train of raw EMG signals is presented in the upper panel of Figure 11.6. As noted in the preceding, some filtering of the raw EMG signal is performed to increase the signal-to-noise ratio, decrease 60-Hz or EKG/EEG artifact, and reduce intersite cross-talk. The primary energy in the surface EMG signal lies between approximately 10 and 200 Hz (Hayes, 1960; van Boxtel, Goudswaard, & Shomaker, 1984). Between 10 and 30 Hz, this power is due primarily to the firing rates of motor units; and beyond 30 Hz this is due to the shapes of the aggregated motor unit action potentials (Lindstrom, 1970; Petrosky & Phillips, 1982). Attenuating the high frequencies in the EMG signal (e.g., using 500-Hz low-pass filters) reduces amplifier noise but rounds peaks of the detected motor unit action potentials. Retaining sharp signal peaks may be important for waveform or spectral analysis of signal motor units or motor units action potential trains but is less critical for obtaining overall estimates of muscle tension. Attenuating the low frequencies (e.g., using 70-Hz high-pass filters) reduces 60-Hz noise from AC power lines, EEG and EKG artifacts, and to some extent, intersite cross-talk (due to the intervening tissue's preferential transmission of low frequencies) but also eliminates a significant portion of the EMG signal. Use of an overly restricted EMG signal passband may result in inaccurate appraisal of the level and form of EMG activity. Hence, selection of an EMG detection passband must proceed based on susceptibility to artifact, presence of extraneous electrical noise at the source and high-frequency noise internal to the amplifier, consideration of the amplitude of the EMG signals to be detected, need to minimize cross-talk, and variations across conditions in muscular fatigue. A passband from 10 to 500 or 1000 Hz is satisfactory for wide-band monitoring; if low-frequency artifact and intersite cross-talk are problematic, then a 90- or 100-Hz high-pass filter may be used, but the investigator should realize one consequence of this selection is that weak signals from the target region will be attenuated as well.⁵

Frequency (e.g., Fast Fourier Transform, or FFT) analyses have occasionally been performed on surface EMG recordings to determine whether there are shifts in the

The skeletomotor system

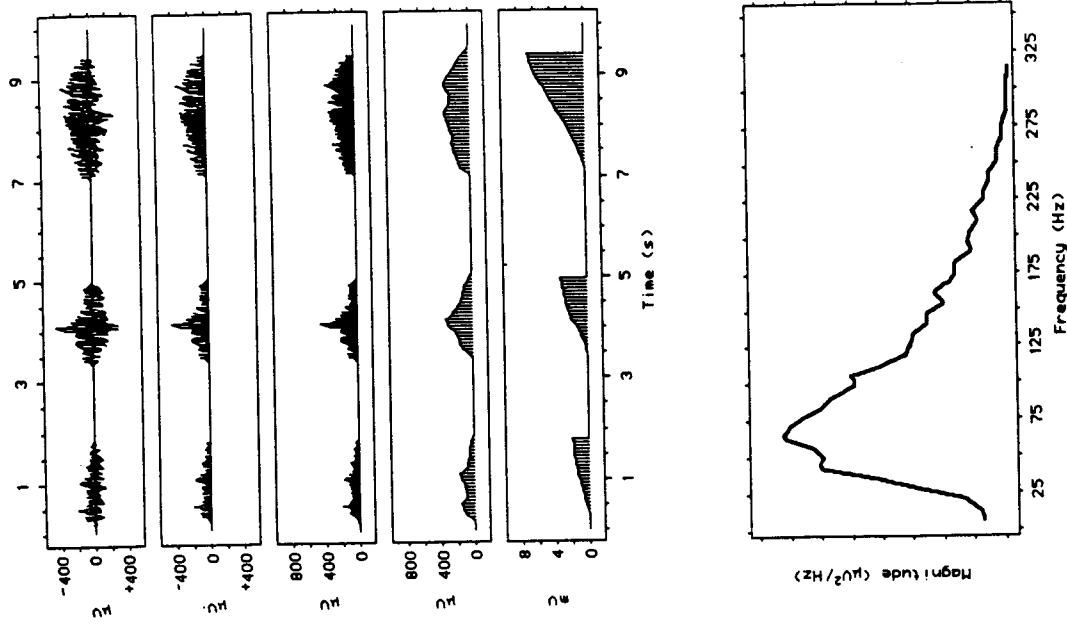


Figure 11.6. Schematic representations of EMG signals. From top to bottom: Raw EMG signal, full-wave rectified EMG signal, full-wave rectified and smoothed EMG signal, full-wave rectified and integrated EMG signal, and EMG power spectrum.

EMG spectra (i.e., magnitude or power at different frequencies) as a function of some psychological or behavioral variable. One robust finding is that shifts in EMG spectra are observed with fatigue (e.g., Mulder & Hulstijn, 1984). Part of the lack of attention to spectral analyses of the surface EMG in psychophysiology is attributable to the following factors: (1) to perform spectral analyses requires that the raw EMG signal be recorded at high rates (see what follows), making the application of FFTs costly; (2) unlike respiration or heart rate, the EMG is not an inherently periodic

physiological event but rather is semistochastic; and (3) thus far, spectral analysis has proven no more sensitive to the few psychological manipulations that have been examined than the relatively inexpensive process of rectification and smoothing (e.g., see McGuigan, Dollins, Pierce, Lusebrink, & Corus, 1982).

The terms *integration* and *smoothing* are often confused. Integration is the temporal summation or accumulation of EMG activity, whereas smoothing refers to integration with a built-in signal decay and is accomplished by low-pass filtering, envelope detection, or averaging of the signal (see Figure 11.6). The total energy in an EMG spectra at any moment in time is equivalent to the rectified and smoothed EMG response; hence, when frequency components of the raw signal are not of interest, considerable economy in terms of data acquisition and signal processing can be achieved by rectification and smoothing prior to digitization and data storage.

The most frequently used on-line "smoother" in psychophysiological research is the contour follower, which is a precision rectifier connected to a low-pass filter (Fridlund, 1979). It acts as a running averager of ongoing EMG activity by providing a varying voltage proportional to the envelope of the EMG signal. Precision contour followers usually have selectable time constants. Short, in contrast to long, time constants produce sensitivity to momentary fluctuations in EMG signals. When measuring a rapidly changing EMG signal, a long time constant blurs rapid changes and signal peaks, and the output can therefore underestimate or overestimate EMG signal strength at any point in time. Conversely, if EMG signals that vary slowly are of primary interest, short time constants will be too sensitive to momentary EMG fluctuations, and the economic advantages of smoothing will be sacrificed.

11.3.2.3 Signal recording, digitization, and storage

Output from preamplification and preliminary signal processing may then be passed to power amplifiers to drive such devices as pen-writing galvanometers, FM tape recorders, and analog-to-digital (A/D) converters (see Figure 11.4). Characteristics of the signal passed from preamplification to this stage determine decisions made at the amplification stage. For instance, the gain required in amplification is dependent on the amplitude characteristics of the signal passed from the preamplifiers.

The rate at which the analog output of the amplification stage is digitized (i.e., the sampling rate for A/D conversion) is a function of the highest frequency present in the signal passed through the amplification stage. The Nyquist relation specifies that signals be sampled at a rate at least twice that of the fastest EMG frequency component of interest. The Nyquist relation, however, assumes both perfect filtering and continuous sampling. A good rule of thumb in laboratory practice, therefore, is to sample at four to eight times the highest frequency of interest to avoid aliasing and allow reconstruction of the original waveform with minimal smoothing. For instance, sampling rates of several kilohertz per channel might be required when digitizing the raw EMG signal, whereas sampling rates of 10–200 Hz per channel can yield satisfactory outputs if precision contour followers with moderate to long time constants are applied prior to digitization.

The process of digitization is accomplished by devices known as A/D converters, and their precision and noise characteristics are important to consider as well. For example, surface EMG recording over the *corrugator supercilii* muscle region can yield signals from fractions of a microvolt to several hundred microvolts. The standard 8-bit A/D converter, although widely available and inexpensive, offers

insufficient resolution (1 part in 256) for detecting weak changes in EMG activity across a range in excess of about 100 μ V. Twelve and 16-bit A/D converters offer a resolution of 1 part in 4,096 and 65,536, respectively; they consequently allow both very weak and very strong muscular contractions to be digitized precisely. One drawback to more precise A/D converters is that they require longer conversion times, placing a ceiling on sampling rates. Fast and precise 12–16-bit A/D converters are now widely available with conversion rates in excess of 100 kHz.

Most A/D converters offer input multiplexers that can be configured for single-ended or differential input. One might assume that passing each EMG channel from the differential amplifier to a single input on the A/D converter would be adequate since the output of the amplifier is the signal of interest. There is a potentially serious problem with this procedure, however. Any variations between the local ground at the amplifier and the local ground for the A/D converter, no matter how minute, can confound detection of the EMG signals being digitized. To avoid this possible source of noise, the input to the A/D converters can be configured for differential recording, and each output from an EMG amplifier can be "differenced" from the ground at the amplifier.

Data storage can be accomplished in analog or digital form. Chart and pen recordings offer a simple visual aid in identifying likely artifacts, but high-frequency signals (e.g., EMG > 75 Hz) are truncated, and digitization of these chart and pen recordings is inefficient and potentially unreliable. Oscilloscopes allow on-line visual inspection of the full range of EMG frequencies, but they are not mass storage devices. Frequency-modulated or pulse-code-modulated tape decks are often useful because they can store the original signal indefinitely with minimal distortion of its amplitudes or frequency components, but decks with a wide recording bandwidth are expensive. Assuming the raw EMG signal was recorded, EMG activity can be replayed either through contour followers with varying time constants and various filters or directly to the A/D converters for purposes of comparison or analysis. Alternatively, the EMG signal can be digitized on-line and stored in digital form. This alternative can be satisfactory if forethought is given to matters such as on-line signal conditioning (e.g., filtering) and A/D conversion rate.

Finally, to ensure that the digitized signals are, in fact, in reasonable correspondence with the input signals emanating from the subject, a precision signal generator can be used to simulate low-level biological signals. These known signals should be varied from zero to several hundred microvolts and from 10–1,000 Hz to mimic the range of EMG signals. Problems such as extraneous electrical interference; amplifier or computer bus noise; improper preamplifier, amplifier, or A/D converter calibration; and undersampling (and consequent aliasing) manifest as discrepancies between the (known) input and digitized signals. This procedure should be repeated for each channel periodically and each time an addition or modification is made to the signal processing sequence. Illustrative designs for signal generators are available in Sheatz (1972) and Helmer (1986).

11.3.2.4 EMG signal representation

Electromyographic activity unfolds over time, and like most other psychophysiological responses, the raw signal is too complex to analyze without considerable data reduction. Whether features of the EMG signal are to be represented in the time domain, amplitude domain, or frequency domain, the first step involves the

conversion of the digitized signal to a descriptive (e.g., physiological) unit of measurement. The numbers assigned to EMG signals of different amplitudes depend on: (1) the electrical unit chosen for description of the signal, (2) the accuracy of the calibration procedure and amplifier's gain setting, and (3) the type of integration method and length of time constant or reset criterion used. We focus on the former factors here. Interested readers can consult Fridlund and Cacioppo (1986) and Basmajian and DeLuca (1985) for additional details.

EMG activity as a voltage-time function. Reference is often made to digitizing analog bioelectrical signals at each peak (e.g., see Cooper, Osselson, & Shaw, 1980, p. 237). In laboratory practice, however, analog signals more commonly are digitized at equal time intervals of duration $1/(\text{sampling rate})$ (e.g., once every A/D conversion). Hence, EMG signals can be viewed as a voltage-time function, where: (1) the ordinate represents bounded signal amplitudes scaled typically in terms of microvolts (μV) and (2) the abscissa represents discrete intervals of time of width $1/(\text{sampling rate})$. The quantification of the amplitudes at each recording/conversion cycle is determined by the direction and magnitude of the measured voltage and is expressed in units of microvolts (μV) or microvolts peak-to-peak (μV_{p-p}). (The latter terminology is reserved for measurements of amplitudes made from negative to positive peaks). The EMG voltage-time envelope, like the motor unit action potential, is bipolar and asymmetrical about electrical zero.

Most psychophysiological research using EMG has focused on some variation of EMG signal amplitude as the dependent variable (cf. Cacioppo, Marshall-Goodell, & Dorfman, 1983). Recording EMG signals with AC-coupled amplifiers for DC amplifiers with zero offsets) ensures that the average value will be zero. Hence, simple averaging of the raw EMG amplitudes is uninformative.

Counting or averaging the EMG signal's peaks, or tallying its directional changes or zero crossings, are relatively easy methods to implement and are useful for gauging gross differences in EMG activity provided a sufficiently high sampling rate is used (Grieve & Cavanaugh, 1973; Willison, 1963). These parameters are further limited in that they do not vary linearly with contraction and are unreliable at the extremes of muscular contraction.

Lippold (1967) maintained that the total energy in an EMG signal at a given moment in time, or what Lippold referred to as the integrated EMG (IEMG) signal, represents overall muscular contraction more accurately than the number or average amplitude of peaks in the EMG signal. Subsequent research has largely corroborated Lippold's assertion given limb movements are constrained and contractions are not extreme (Basmajian & DeLuca, 1985; Goldstein, 1972; Winter, Rau, Kadefors, Broman, & DeLuca, 1980). However, the term IEMG has been used in this research to refer to several different processing techniques. Two of the most common processing techniques in contemporary research are (1) the rectified EMG signal, smoothed using a contour follower, expressed as average voltage (μV or μV avg), and (2) the root-mean-square (rms) EMG, which is expressed as an rms voltage (μV or μV rms).⁶ Both processing techniques transform the EMG voltage-time function into a contoured waveform that is nonnegative and bounded in time and amplitude. The moment-by-moment amplitude of this function represents an estimate of the total energy of the signal at the time; the mean amplitude of this voltage-time function represents the average level of electrical energy emanating from the underlying muscle region(s) during a given recording epoch; and the

integral of this function (e.g., the sum of the amplitudes) represents the total electrical activity (i.e., the size of the response) emanating from the underlying muscle region(s) during the recording epoch. Mean amplitude and total electrical energy are redundant when the sampling rate and length of the recording epochs are constant across conditions.

Cacioppo, Marshall-Goodell, and Dorfman (1983) demonstrated that although traditional measures of smoothed or integrated EMG amplitudes are generally representative of the level of isometric muscular contraction, they are insensitive to the distribution of amplitudes and to the manner in which the EMG signals emerge across time. Thus, one consequence of the traditional focus on the amplitude domain of the EMG signal is that the form of the response across time has been largely ignored. A notable exception is Malmo's (Davis & Malmo, 1951; Malmo, 1965, 1975) use of *EMG gradients*, which he defined as a "progressively rising level of voltage reflecting gradually rising tension in the specific muscle group on which the electrodes are placed" (Malmo, 1975, p. 34). Electromyographic gradients have been used to assess variations across time in tonic muscle tension and are depicted by plotting the mean (or rms) EMG amplitude for two or more consecutive recording epochs.

Cacioppo and Dorfman (1987) (see also Dorfman & Cacioppo, chapter 20) have recently extended work on representing EMG waveforms in the time, amplitude, and frequency domains to derive a systematic and comprehensive representation of nonnegative bounded waveforms predicated on the moments of the waveform in a given domain. Nonnegative bounded waveforms are an important class in psychophysiology, ranging from IEMG activity in the time domain, EMG or IEMG activity in the amplitude domain, and the power spectrum of the EMG signals. The waveform of EMG signals can be distorted in numerous ways, including threshold settings, rectification, and filtering (Loeb & Gans, 1986). Although an important goal in EMG recording is to preserve the original signals with as little distortion as possible, a more serious problem for waveform analyses is when distortion in a signal (e.g., due to filter settings) varies across comparison conditions.

Baselines. As in psychophysiology generally, it is often desirable to obtain both basal measures of activity and response measures that are uncontaminated by the basal (e.g., prestimulus) level of activity. The notion of "basal" activity can be ambiguous when applied to EMG signals, however. This is because the true "physiological baseline" for EMG activity is zero; hence, the lowest empirical baseline for EMG recording is the system noise floor. On these grounds, Fridlund and Izard (1983) questioned the pro forma baseline correction of EMG scores.

In laboratory practice, EMG recording levels seldom show zero activity because the alert experimental subject is rarely completely relaxed. Therefore, it is important to consider the EMG activity that exists in the absence of experimental stimuli both to assess individual differences and to help achieve a measure of the experimental treatments free from prestimulus EMG activity. In assessing basal EMG activity, care is required to avoid any confounding of measurements with task-irrelevant activity (e.g., adaptation, fatigue, apprehension). The procedures commonly used include recording during prestimulus periods and recordings during pseudotrials (Johnson & Lubin, 1972). The use of pseudotrials has the

advantage that assessments are obtained under conditions identical to the experimental trials except that there are no experimental stimuli to which to respond.

Obtaining measures of the response free from the influence of prestimulus levels can be more complicated. Perhaps the oldest procedure is to subtract the prestimulus period mean from that observed during the stimulus period. This procedure is satisfactory if mean amplitude is the only parameter of interest and the prestimulus state will not confound the stimulus period state (cf. Cacioppo, Marshall-Goodell, & Dorfman, 1983, pp. 276-277). On the other hand, significant differences across conditions in prestimulus period EMG activity contraindicates the use of this procedure.

A closed-loop baseline procedure offers an alternative to the use of simple change or residualized scores (McHugo & Lanzetta, 1983). Briefly, the presentation of experimental stimuli or treatments is programmed to be contingent on acceptably low levels of somatic activity across the recording sites. In this way, task-specific EMG responses are quantified while minimizing the confounding effects of extraneous muscular activity or basal differences in somatic activity across treatments, a procedure that is reminiscent of Jacobson's (1932) use of progressive relaxation in studies of EMG and imagery. Using a closed-loop baseline has the advantage over simple change scores, too, in that time series (see Gottman, chapter 22) and waveform moment analyses (see Dorfman & Cacioppo, chapter 20) can be performed with fewer restrictions. A potential liability of the closed-loop procedure is that achieving low levels of EMG activity may be one outcome of a subject's idiosyncratic strategies, and experimenters using the closed-loop baseline may be shaping subjects inadvertently for these strategies.

In summary, because the physiological baseline for EMG signals is zero, the rationale for baseline corrections of EMG signals (e.g., covariance analyses, change scores, range corrections) is weak at best. Fridlund, Schwartz, and Fowler (1984) compared baseline-corrected and range-corrected with uncorrected EMG signals and found that corrected EMG scores did not improve the ability to discriminate facial displays within subjects. It is simply unknown at this time whether these corrections would prove useful empirically in discriminating among treatments between subjects.

Standard scores. Standard scores, range-corrected scores, and change scores have been used in an attempt to achieve a metric common to all sites or subjects. For example, standardizing EMG scores within sites and subjects has been used in an attempt to reduce individual and site variability. Although this procedure can provide a powerful means of attenuating individual variability in EMG research, it has limitations that are often unrecognized. Dorfman (cited in Fridlund & Cacioppo, 1986) offered the following numeric illustration (see Table 11.1).

Briefly, let T_1 represent a within-subjects treatment and S_j represent a particular subject. Mean EMG measures are presented in the left columns of Table 11.1, and the corresponding standardized EMG measures are in the other two sections. Even though standardization was within subjects, the order of the treatment means T_1, \dots, T_5 , and the intervals between individual data points within subjects were altered by standardization. If one had used only T_1 and T_2 in this study, both untransformed and z-scores would have indicated that $T_2 > T_1$ (see middle columns of Table 11.1). Such a result might obtain, for instance, if EMG activity over the *corrugator supercilii*

The skeletomotor system

Table 11.1. Hypothetical data showing hazards of EMG score standardization

Subjects	Raw scores in hypothetical study 1 ($T_1 - T_2$) and study 2 ($T_1 - T_2$)					Standard scores in hypothetical study 1					Standard scores in hypothetical study 2				
	T_1	T_2	T_3	X	SD	T_1	T_2	T_3	X	SD	T_1	T_2	T_3	X	SD
	S_1	3	2	1	2.00	1.00	0.707	-0.707	0.0	1.00	1.00	1	0	-1	0.00
S_2	1	2	10	4.33	4.93	-0.707	0.707	0.0	1.00	1.00	-0.675	-0.473	1.15	0.00	1.00
S_3	1	2	10	4.33	4.93	-0.707	0.707	0.0	1.00	1.00	-0.675	-0.473	1.15	0.00	1.00
X	1.67	2.00	7.00			-0.236	0.236				-0.120	-0.320	0.43		

muscle region were being recorded and T_1 and T_2 represented imagery conditions that generally evoked mild happiness and anger, respectively.

If this within-subjects study were then perfectly replicated (i.e., the same raw data were obtained in response to T_1 and T_2), but treatment T_3 was added (where, e.g., T_3 represented imagery conditions that generally evoked sadness), then the untransformed scores for T_1 and T_2 would, by definition, remain unchanged. Such a result would, of course, be interpreted as being consistent with the previous study showing that the induction of sadness or anger is associated with increased EMG activity over the *corrugator supercilii* muscle region. Inspection of the raw data reveals there is considerable individual variability in EMG activity, however, and one might reason that these data should be standardized within subjects, purportedly to obtain a measure of the effects of treatments free from individual differences.

Inspection of the z-scores in the right columns of Table 11.1 reveals the potential flaw in this reasoning; the standardized EMG measures now make it appear that $T_1 > T_2$. If the raw data were not first inspected, such a result would be interpreted as a failure to replicate the previous study even though the actual outcomes of hypothetical studies 1 and 2 were identical. This apparent failure to replicate is attributable entirely to an erroneously applied nonlinear transformation. The point is not that EMG data should never be standardized or transformed, but rather that for the sake of replicability, the application of any nonlinear transformation of EMG signals for psychophysiological data should be accompanied both by an explicit justification (cf. Levey, 1980) and an acknowledgment of any differences in the ordering of the untransformed versus transformed means.⁷

11.3.2.5 Statistical analysis

Perhaps due to the large number of muscles and to the fact that their coordination is required to communicate, anticipate, adapt, and react, it is analyses of the form or pattern of somatic activity - across sites and across time - that are potentially most informative in psychophysiological research. Before proceeding to a review of some of this research, we discuss briefly selected issues involved in statistical analyses of these data. A more detailed discussion of multivariate analyses of psychophysiological data is provided by Russell (chapter 23; see also Fridlund & Izard, 1983).

Multivariate procedures have several advantages over univariate analyses in EMG research. They provide a quantitative index of whether or not the obtained measures (or subsets of measures) varied as a function of conditions. Second, they can provide information about the configuration of responding across time (e.g., Cacioppo, Petty, & Marshall-Goodell, 1984) or sites (e.g., Fridlund et al., 1984).

There are problems in relying entirely on multivariate procedures, however. First, research in which multiple psychophysiological measures are obtained tends to be expensive to conduct and can yield huge quantities of data even though the number of cases (e.g., subjects) is relatively small. Most multivariate procedures, however, require 8-10 cases per variable per treatment group. An insufficient number of cases in the sample threatens the reliability of the configurations (factor weightings) identified in multivariate analyses of these data.

Second, the results of the analyses can depend heavily on the particular set of variables (in particular, their variance-covariance structure) included in the analysis (Harris, 1975). This is because the inclusion or deletion of another psychophysiological measure or set of measures can affect the variance-covariance structure of the data and lead to dramatic changes in the weighting (i.e., apparent patterning) of one or more measures obtained in the original multivariate analysis. Instability in weights in a multivariate analysis is less of a problem when (1) the psychophysiological measures are not highly correlated, a situation which lessens the chances of variables serving a suppressor or moderator influence on the solutions; (2) obtaining a multivariate effect is sufficient without regard for the dependent variable weights that accounted for it, as when using MANOVA to assess whether or not there is an overall difference across treatments; (3) cross-validation on new data establishes the robustness of the multivariate solution; and (4) the directions of the univariate effects corroborate the multivariate solutions (Fridlund & Cacioppo, 1986). Multivariate data analyses, which appear to reveal order not discernible in univariate or graphical analyses, should be interpreted cautiously, however.

11.3.3 Social recording basis

Social factors and processes are of increasing interest to psychophysiological researchers. First, psychophysiological research was once thought to be exempt from the laboratory artifacts that have led others to consider the physical and social context in which the research was conducted (e.g., Rosenthal & Rosnow, 1969). The vulnerability of physiological responses to instructional sets (Sternbach, 1966), intentional distortions (e.g., Ekman, 1985; Honts & Hodes, 1982), and social biases (Ekman & Friesen, 1975; Tognacci & Cook, 1975) vitiates this notion. Nowhere is this vulnerability more apparent in psychophysiology than in studies of the skeletomotor system (Cacioppo, Petty, & Tassinari, 1989; Fridlund & Izard, 1983). A serious consideration of these factors can contribute to the construction of more sensitive, artifact-free psychophysiological experimentation and, hence, to stronger inferences. We begin this section by briefly discussing two major ways in which social factors can contribute to artifactual results in EMG research. Interested readers may also wish to consult Cacioppo, Petty, and Marshall-Goodell (1984, 1985), Christie and Todd (1975), Gale and Baker (1981), Fridlund and Izard (1983), Christie and (1987), and Fridlund and Cacioppo (1986).

A second reason for situational and social factors gaining wider attention in psychophysiology is that these factors have been found to moderate the influence of nonsocial factors on physiological responding. This point, too, is perhaps clearest in studies of skeletomotor response. The expression of a person's joy following successful completion of a task, for instance, can be magnified, attenuated, or

The skeletomotor system

masked because of the presence of others. Comprehensive psychophysiological theories must accommodate such moderating influences. We therefore end this section with a brief survey of social moderators of EMG responses to nonsocial stimuli or states. Interested readers may wish to consult Cacioppo, Losch, et al. (1986) and Fridlund (in press) for additional discussions of these and related issues.

11.3.3.1 Social factors as laboratory artifacts

As Gale and Baker (1981) noted: "In psychophysiological studies, experimenter-subject interactions are particularly important since the procedures may involve bodily contact, partial removal of clothing, skin abrasion, touching, and application and removal of electrodes" (p. 373). This somewhat unique and extended interaction between subjects and experimenters may result in subjects becoming anxious, distracted, or aware of the experimenter's expectations. In any case, significant laboratory artifacts may be introduced.

Demand characteristics. Orne (1962) suggested that many subjects try to discern the true purpose of an experiment and behave accordingly. To the extent that subjects can discern the experimental hypotheses and manipulate their actions accordingly, EMG studies are vulnerable to experimental demand confounds. Fridlund and Izard (1983) argued that many facial EMG studies of emotion and emotional imagery could be interpreted in terms of demand characteristics. Specifically, they reasoned that placing multiple electrodes on a person's face can make the person acutely aware of her or his facial expressions. Subjects who desired to please the experimenter or contribute to "science" might make faces in accordance with what they believed were the experimental hypotheses. Particularly susceptible, Fridlund and Izard (1983) suggested, were studies in which the experimental hypothesis was obvious, such as the early studies of facial EMG activity during emotional imagery (cf. Schwartz, 1975). Not all previous research using facial EMG to study emotion was susceptible to this interpretation (cf. Cacioppo & Petty, 1979a), and recent studies whose designs minimized experimental demands have still found that subtle emotions could be discriminated using facial EMG (Cacioppo et al., 1984; Cacioppo, Petty, et al., 1986; Fridlund et al., 1984; McHugo, Lanzetta, Sullivan, Masters, & Englis, 1985). Hence, experimental demands do not appear to be necessary for facial EMG patterning during emotions, but they nevertheless remain a potential source of bias in many EMG studies.

A variety of procedures can be used to minimize or assess the effects of experimental demands in EMG research. These include: (1) designing the laboratory setting to minimize subjects' perceptions of being scrutinized; (2) giving subjects peripheral experimental hypotheses; (3) using dummy electrode placements over bodily areas that lower subjects' awareness that voluntary bodily responses are being monitored; (4) employing cover stories that divert subjects' awareness from the fact that physiological actions over which they could exert voluntary control are being recorded; (5) challenging subjects to guess the hypothesis at the end of the study and analyzing suspicious subjects separately; and (6) conducting "simulation" studies in which the EMG study is described to a naive group of subjects, who are then asked to predict how subjects would respond in the setting or how they should respond to confirm the experimenter's hypothesis.