

## Subjective Symptomatology During Work and Fatigue

ROBERT A. KINSMAN and PHILIP C. WEISER

THE CONCERN of this chapter centers around what has been called "subjective fatigue"—the feeling states experienced during the performance of work. Changes in these feeling states often may be related to work tolerance and performance. For example, individuals engaged in sports or any type of physical activity involving prolonged, strenuous exercise must continually determine their level of energy expenditure and the length of time they are willing to continue work. In general, the point at which a person quits may correspond to when he "feels very tired" or "worn out."

Welford (1965), however, has noted that the symptoms commonly associated with subjective fatigue have been viewed as difficult to measure adequately. Consequently, until recently they have been neglected by most investigators in favor of more directly observable physiological, neurological, or behavioral events. Indeed, a search of the literature in this area quickly reveals many inadequacies in methodology and a paucity of well directed and systematic investigations of subjective symptomatology during work. Fortunately, as will be seen, there are also some promising experimental approaches which have recently appeared in the literature.

It is certainly true that subjective symptoms cannot be measured directly. By definition, a symptom is a privately experienced event made public only when the experience is reported to an observer. The measurement of these symptoms is, therefore, necessarily indirect, depending on reliable self-report techniques. Furthermore, it is important to realize that the subjective symptoms most of us usually associate with fatigue *per se*, such as tiredness, lack of energy, or feeling worn out, probably encompass only a few of the many types of subjective symptoms that might be reported during work performance. Other subjective qualities as diverse as pain, thermal sensations, or boredom often occur during work. It is important to differentiate between these subjective qualities as Welford (1965) has pointed out. In recent years, there has been renewed effort to develop self-report techniques to measure such subjective qualities accompanying work performance.

It has frequently been assumed that subjective symptoms must inevitably

be related to work performance under any task condition. Some authors have suggested that the opposite is true, that is, that subjective symptoms are not related to work performance (e.g. Pierson, 1963a). As with many extreme positions in science, both statements are overgeneralizations. Frequently there have been work situations where no relationship between symptomatology and performance have been reported or would be expected. While this lack of relationship could result from the selection of the method of measurement or the specific quality measured, it may adequately reflect the real situation.

The objectives of this chapter are first to review and evaluate studies of the various subjective qualities reported during work, and second, to propose a model permitting a more appropriate conceptualization of subjective fatigue. The review section, comprising roughly the first three-fourths of the chapter, is not exhaustive but serves to exemplify the methods of measurement used and the subjective qualities studied. The reader is forewarned that the review often involves detailed discussion of individual studies since an adequate, critical overview of this area does not appear to exist. We hope the persistence of the interested reader will be rewarded by the novel conceptual organization of the area provided in the last quarter of the chapter.

### THE MEASUREMENT OF SUBJECTIVE QUALITIES

Subjective symptoms represent one level of measurement during work performance in a matrix (the organism) which consists of a diversity of organizational levels. Questions should be asked about relationships within as well as between levels of biological organization during work performance: What are the subjective symptoms experienced during work? How do these differ between tasks? How do these change during the course of work? What are their relationships to physiological, and especially neurological, levels of organization? When are they related to work performance and work tolerance? When are they not related? None of these questions presume a specific answer, but the answers are bound to be informationally rich. If the first order of business in studies of subjective symptoms during work—measurement—is handled adequately, the answers to such questions as these will be interesting in their own right.

### Scaling Techniques

There are a variety of solutions to the problem of adequate measurement that are available from the areas of psychophysics (e.g. see S. S. Stevens, 1957), psychological assessment (e.g. see Jackson and Messick, 1967), and, in particular, psychometric scaling (e.g. Guilford, 1954). Recent developments in scaling techniques have been reviewed by Cliff (1973) in an ar-

ticle containing over 200 references for the period 1968 to 1972 alone. The available scaling techniques have been used only recently in a *systematic* program of research in the area of work performance.

S. S. Stevens (1960) has discussed the numerical properties of various types of scales. Before proceeding further, it may be useful to review the characteristics of scales. A scale simply refers to the assignment of numbers to objects or events in order to differentiate between them. *Nominal* scales are the simplest means of numerical assignment providing only a substitute name for the purposes of identification. Social Security numbers, automobile license plates, and the numbers given to football players are all examples of nominal scaling. No information is provided about the magnitude of some quality of the objects to which numbers are assigned. The automobile with a license plate AI-4516 may either be a Volkswagen sedan or a Cadillac Eldorado according to nominal scaling.

In an *ordinal* scale, numbers are assigned according to the order of magnitude for some attribute which objects or events possess. For example, a compulsive psychometrician may ask his wife to rate how hungry she is on a 5-point scale before deciding on the restaurant for dinner. He may say "suppose a 5 is the hungriest you've ever been and 1 is the least hungry you've ever been. How hungry are you on a scale of 1 to 5?" If she's willing to put up with this nonsense, she may reply 1 indicating that she's very hungry, but not as "hungry as she's ever been." The scale thus allows a rough quantification of a subjective quality, hunger, and the psychometrician can use this information in making his choice of restaurants for dinner. In such an ordinal scale, however, there is no true zero point representing the complete absence of hunger, nor are the intervals on the scale necessarily equidistant. The difference between 4 and 5 or 1 and 2 on the "hunger scale" may both be 1 scale point, but may not represent equal differences in the magnitude of the subjective quality measured. Nevertheless, most applications of scaling techniques to measure subjective feelings during work performance have used rating scales with ordinal properties. As we will see, the shortcomings of ordinal scales, absence of a "true" zero point and equal intervals, can largely be overcome in most applications.

*Interval* scales possess all the attributes of ordinal scales, but the intervals between adjacent categories are equal. A thermometer is the most familiar interval scale with degrees indicating the magnitude of the difference between points on the scale. However, like ordinal scales, the interval scale lacks a true zero point. The zero on a Centigrade or Fahrenheit thermometer does *not* correspond to the "absence of heat." Interval scales can be constructed to measure subjective qualities, but require the use of rather lengthy psychometric techniques such as the method of equal appearing intervals (Thurstone and Chave, 1929; also see Niven, 1953; Guilford, 1950.

1954). In the measurement of subjective qualities during work performance, interval scaling is best exemplified by the studies of Pearson and Byars (1956) and of McNelly (1957), both of which will be discussed in some detail.

Only the *ratio* scales have a true zero point representing the absence of the attribute measured while also providing magnitude measurement by equal intervals in the units of measurement. Psychophysical techniques are used to obtain ratio scales for subjective qualities.

Most attempts to quantify subjective qualities during work have used *ordinal* rating scales. The inherent shortcomings of the ordinal scales can largely be overcome in most applications. The intervals between ordinal scale points often can be made to *appear* equidistant by careful selection of the verbal descriptions accompanying each scale point, physically representing the verbal descriptions at equal intervals, and by indicating equal intervals in the numbering of scale points, e.g. values of 0, 1, . . . , 4. Furthermore, a zero point can be provided by anchoring the first category with "not present" indicating complete absence of the subjective quality. Such ordinal scales have often been treated statistically as if they were true equal-interval scales. Overall and Klett (1972) have noted that equal-appearing interval scales "have frequently been demonstrated to be linearly related to Thurstone type equal-interval scales except in the extreme categories." Such scales are among the simplest to construct and to use; therefore, they have enjoyed the widest application in studies of subjective fatigue to be discussed.

#### TYPES OF MEASUREMENT OF SUBJECTIVE QUALITIES

In this section the techniques available to measure subjective qualities during work performance will be reviewed and evaluated. Measurement has relied primarily on two general methodologies: (1) psychophysical techniques and (2) rating scales. Psychophysical techniques produce ratio scales which although more precise in describing relationships between stimulus levels and perceptions, are less useful in most work situations. Rating scales have enjoyed a wider application largely because of the ease with which the ordinal or interval scales may be used. Most recently, a third general methodology involving multivariate statistical procedures has been used to identify categories or dimensions of symptomatology related to work performance.

##### Psychophysical Studies

There has been renewed interest in the application of psychophysical techniques to relate subjective qualities to certain parameters of work performance (see Table I-1). These techniques provide a *ratio scaling* that relates stimulus level to the perception of the stimulus. Application of psy-

chophysics to the area of work performance provides rather precise answers to questions regarding how the perception of effort increases as physical activity increases.

Classical psychophysics has a long history extending back to the 19th century and the work of Weber (1834), Fechner (1877), and Plateau (1872). Detailed discussions of the development of psychophysics have been presented by S. S. Stevens (1957, 1960). Psychophysics involves the description of the relationship between *perceived* change in a physical stimulus, the psychological dimension, and change in the objective intensity or quality of the stimulus measured in physical units, the physical dimension. For many years, it was assumed that Fechner had accurately described the relationship between perception and levels of a physical stimulus as a logarithmic function:

$$\Psi = k \log \Phi \quad (1)$$

where  $\Psi$  is the subjective unit of the psychological scale,  $k$  is a constant and  $\Phi$  is the level of the physical stimulus.

Since the 1930's, however, under the influence of modern psychophysicists such as S. S. Stevens and Ekman, this relationship is generally agreed to be better approximated by a power function:

$$\Psi = k (\Phi)^m \quad (2)$$

where  $m$  is an exponent defining the power function fitting the shape of the curve.

The value of the constant threshold, i.e. the minimum stimulus level which the individual can perceive may also be included in the power function as a constant:

$$\Psi = k (\Phi - a)^m \quad (3)$$

where  $a$  is an additive constant included to bring the zero point of the physical scale in line with the zero of the psychological scale. Inclusion of this constant in the equation will generally result in a rescaling of stimulus values corresponding to the range of perception. According to S. S. Stevens (1960), it was Luce (1959) who noted that "the use of an additive constant to bring the zero of the physical scale into coincidence with the zero of the psychological scale is a proper generalization of the power law."

This power function has been found to be applicable to subjective judgments covering a wide range of stimulus modalities including brightness, pitch, loudness, taste, odor, and heaviness (S. S. Stevens, 1957) as well as the perception of force exerted during work performance (Borg, 1962; Schmale, Schmidtke, and Vukovich, 1963; and Bernyer, 1967), and fatigue (Bujas, Pavlina, Sremec, Vidaček, and Vodanović, 1966). An example of a power function relating perception of exertion ("Anstrengung") to load level expressed as percentage of maximum load is shown in Figure 14-1.

Various psychophysical methods are used to derive the power functions. These have been described in detail in numerous sources (e.g. see Stevens, 1957, 1960; and Ekman, 1958, 1959). It is beyond the scope of this chapter to discuss in detail the many issues and methods in the area of psychophysics. However, it may be worthwhile to describe two of the more useful psychophysical methods of ratio scaling: the *estimation* and *production* methods, both of which have been applied to the perception of force exerted during work.

In general, *estimation* methods require the subject either to estimate the relative magnitude (magnitude estimation) of physical stimuli by assigning numerical values (e.g. 105, 120, etc.) to a series of comparative stimuli in reference to a standard stimulus of a given numerical magnitude (e.g. 100); or, to estimate the percentage magnitude (ratio estimation) of the comparative stimuli in relation to the standard. The *production* methods, in contrast, require the subject to *produce* a stimulus that is either a multiple of (ratio production) or a specified magnitude (magnitude production) of a standard stimulus of given intensity. In work situations, the stimulus dimension in a psychophysical study might typically be work load (e.g. kpm/min. on ergometer tasks), weight lifted, and grade or walking speed on a treadmill. By these psychophysical methods, ratio units of perceptual values may be obtained based on the judgments associated with the sequence of comparative stimuli presented in relation to the standard stimuli used. The psychophysical relationship between perceptual values and values of the physical stimuli can frequently be described by equation 2 or 3 above. When this function is plotted on log-log coordinates, the exponent of the power function ( $m$ ) defines the slope of a straight line, permitting direct comparisons between perception-stimulus relationships for different stimulus modalities. In general, a high exponent indicates that the perception of the stimulus changes rapidly with respect to changes in the physical stimulus; in contrast, a low exponent indicates that the perception changes relatively more slowly with respect to the physical dimension. For example, exponents of the power function for loudness have been found to be about .6 indicating that the perception of loudness grows slowly relative to changes in sound intensities. In contrast, the exponent for electric shock has been observed to be 3.5, indicating that perception changes rapidly with respect to shock intensity.

Table 14-1 presents a summary of specific psychophysical studies relating to physical work, indicating the task and psychophysical method used, the subject population studied, the subjective quality measured, and the exponent of the power function obtained. In an initial study, Borg and Dahlstrom (1960) investigated perception of force exerted (pedal resistance) on a bicycle ergometer using four subjects. The ratio estimation technique

TABLE 14-1  
PSYCHOPHYSICAL STUDIES DURING WORK PERFORMANCE

Subjective Quality	Physical Modality	Task	Subjects	Exponent of Power Function	Studies
Force exerted	Work load (kpm min)	Bicycle ergometer	Young males (n = 4)	1.6	See Borg (1962)
Force exerted	Work load (kpm min)	Bicycle ergometer	Young males (n = 12)	1.6	See Borg (1962)
Force exerted	Work load (kpm min)	Bicycle ergometer	Young male students (n = 12)	1.8	See Borg (1962)
Force exerted	Force of pressure (lbs)	A. Stationary foot pedal B. Hand grip	Young males and females (n = 12)	A. 1.63 B. 1.65	Eisler (1962)
Force exerted	Work load (kpm sec)	Arm and leg ergometer	Young males and females (n = 20)	1.4	Schmale, Schmidtko, and Vukovich (1963)
Force exerted	Work load (kpm sec)	Wrist-cutting	Young males and females (n = 6)	1.9	Schmale, Schmidtko, and Vukovich (1963)
Force exerted	Weight (kps)	Weight-holding	Young males and females (n = 10)	3.1	Schmale, Schmidtko, and Vukovich (1963)
Force exerted	Weight (grams)	Lifting weights by manually pressing levers	Young males (n = 3)	1.8 to 2.5	Beyner (1967)
Force exerted	Force of pressure (newtons)	Hand grip	Males (n = 18)	1.7	Stevens and Cain (1970); also see Stevens and Mack (1959) and Cain and Stevens (1971)
General feeling of fatigue	Duration of effort (secs)	Weight-holding	Males (n = 5)	1.5	Bijas, Pauling, Stenec, and Vidachek (1966)

was used with power levels measured in kpm/min. This study resulted in the power function with an exponent of 1.6 and having a high test-retest reliability ( $r = .95$ ). In subsequent pilot work, Borg (1962) determined that length of the work period, varied systematically between 5 and 100 sec, affected the perception of force exerted: longer work periods resulted in higher perceived force exerted for the same task.

Subsequently, twelve male subjects were studied using the ratio production method. Five standard load levels, presented for five seconds, were used: 500, 700, 900, 1,000, and 1,300 kpm/min. On each trial, the subject was required to establish (i.e. produce) a load level equal to one half of the standard. The derived power function had an exponent identical to the earlier preliminary study. An estimate of test-retest reliability for this data was exceptionally high:  $r = .96$ . Perceived force exerted was also studied in a group of twenty-two male subjects using the ratio production method wherein one power level (800 kpm/min.) was halved 10 times by ascending (5 trials) or descending (5 trials) to the perceived mid-point. Calculations indicated that the resulting power function had an exponent between 1.39 and 1.73.

Using ratio estimation, as opposed to the production method described above, Borg (1962) derived the power function for perceived force exerted using twelve male subjects between 20 and 30 years of age. In this study, load-levels of 300, 600, 900, and 1,200 kpm/min. were presented in pairs, one immediately after another, for 20 seconds each. The task for the subject was to estimate the intensity of the lower load level of each pair as a percentage of the higher. Each trial was separated by a 30-second rest period. Six of the twelve pairs were in ascending order, and six in descending order; order of presentation was random. After completion of this experimental sequence, the same subjects were rerun with the pairs presented in ascending order, so that within each pair and across trials, higher stimulus intensities followed lower stimulus intensities. This latter condition was designed to correspond with the stimulus sequence often used in work tests. The obtained power functions had exponents of 1.8 and 1.2 for the random and ascending series, respectively. An ascending series thus appears to lead to a higher absolute threshold for force exerted (300 vs. 200 kpm/min.) and to a less rapid increase in perceived force exerted relative to load level. This effect upon the exponent of the power function has been labeled "hysteresis" (S. S. Stevens, 1957).

Borg's studies suggest that the exponent for perceived force exerted during cycling is 1.1 to 1.8, corresponding quite well to that obtained for other work modalities. For example, perceived force of handgrip was found to have an exponent of 1.7 (J. C. Stevens and Mack, 1959), perceived heaviness of lifted weights has been found to have an exponent of approximately 1.45

using both production and estimation techniques (S. S. Stevens and Galanter, 1957), and the perception of respiratory sensations—pressure, lung volume, and ventilatory volume—have been found to have exponents of 1.5, 1.3, and 1.9, respectively (Bakers and Tenney, 1970).

Eisler (1962) derived a ratio scaling of subjective force exerted for a mixed group of men and women using magnitude production and estimation techniques. The principal interest of this study was to evaluate the exponent of the power function for force exerted when large muscle groups were used. Additionally, Eisler inquired about the correspondence between the exponent for work involving large muscle groups, i.e. exerting a force in a horizontal direction against a foot pedal, and work involving small muscle groups, i.e. squeezing a handgrip. For the foot pedal task, the exponent of the power function for force exerted was found to be approximately 1.6, in excellent agreement with the findings of Borg (1962a) who had used the bicycle ergometer. To compare the exponents of the two work tasks, Eisler used two matching tasks: (a) the subject was required to squeeze a hand grip to match the force previously exerted against the foot pedal, and (b) vice versa. If the perception of force exerted changed with the physical force exerted in the same way for both hand grip and foot pedal tasks, the slope of a plot of force exerted for the hand grip vs. foot pedal matching tasks would equal unity. Results confirmed that, in fact, the slope of both lines (hand grip vs. foot pedal and foot pedal vs. hand grip) were unity. Thus, these results suggest that the perception of force exerted changes with the actual physical force exerted similarly for both large and small muscle groups, with power functions having exponents of approximately 1.6.

Schmale, Schmidtke, and Vukovich (1963) have summarized their own substantial psychophysical studies of work, reporting on a variety of tasks ranging from dynamic work (arm and leg bicycle ergometer tasks, and wrist curling) to static work (holding up a weight). For each of the dynamic work tasks, the production method was used to derive the power function. For the bicycle ergometer tasks, a linear relationship was found between "Anstrengung" (perceived exertion) for both leg and arm work. This linear relationship permitted derivation of a common power function for both arm and leg work relating "Anstrengung" to load level defined in percent of maximum load ( $P$ ):  $\psi = (P + 2.14)^{1.4}$ . The exponent agrees with that reported subsequently by Borg (1962) for dynamic work on the bicycle ergometer. For static work involving holding up weights suspended from the wrist, the power function relating "Anstrengung" to weight expressed as a percent of maximum weight that could be held ( $P$ ) was:  $\psi = .000008 (P + 42.1)^{3.1}$ . A graph of this power function is presented in Figure 14-1. In summary, the exponent of the power function for dynamic

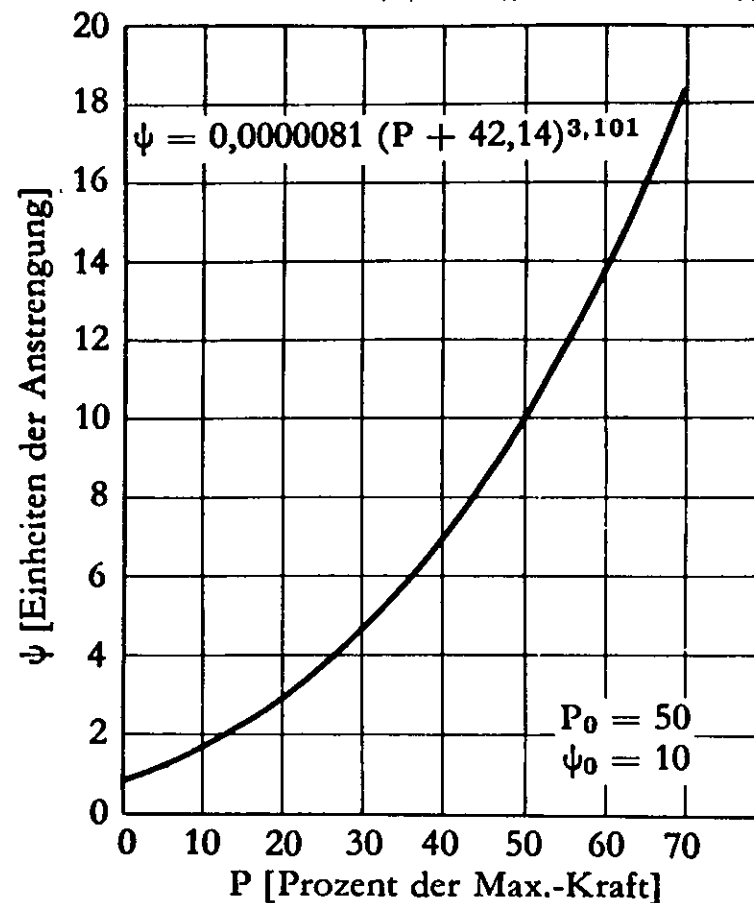


Figure 14-1. Power function relating "Anstrengung" (effort) to weight as a percentage of maximum strength. (From H. Schmale, H. Schmidtke, and A. Vukovich, "Untersuchungen über den Grad der subjektiv gegebenen Beanspruchung bei Körperlicher Arbeit, Forschungsbericht des Landes Nordrhein-Westfalen, Nr 1261, 1963).

work was between 1.4 to 1.9 whether the work involved arms, legs, or wrist curling. However, the exponent obtained for static work, 3.1, was considerably higher.

Bennyer (1967) has demonstrated that production methods (e.g. the method of bisections) may result in different mathematical functions for

perceived heaviness of lifted weights depending on whether the test series presented was in ascending or descending order. For ascending order, a power function with exponents ranging from 1.8 to 2.5 describes the positively accelerated perception-stimulus relationship adequately; for a descending series, a negatively accelerated function is produced, apparently described more adequately by a Fechnerian log function, shown in equation 1 above. The actual relationship derived by psychophysical techniques has been shown to be substantially influenced by task requirements. Among the factors affecting the obtained psychophysical relationships are, as previously noted, the direction of the series presented (e.g. ascending or descending) and time-order errors related to order of presentation of the stimuli in a series (S. S. Stevens, 1957).

Hosman (1967) studied the time-course for the "feeling of exhaustion" in 5 subjects during three prolonged static work tasks. Using cross-modality matching against white noise, 9 of the 15 total responses showed a linear growth function. In 2 trials the increase was negatively accelerated, while in only 4 trials did the time-course show a positive acceleration. However, since the subjective response was reported as an average of 10 sessions at each task, the linear response might partially represent a timing bias that could have developed after several sessions (see p. 370). More detailed and sophisticated psychophysical analysis of the time course for fatigue during prolonged static work has been reported by J. C. Stevens and Cain.

J. C. Stevens and Cain (1970) explored the more complex interaction that the effects of force and duration have upon perceived force of handgrip using 18 male subjects. In an initial study, magnitude estimation was used to obtain a ratio scaling of perceived force relative to handgrip squeezes at forces between 22 and 220 newtons (5 to 50 lbs) for durations of 4 to 60 secs. Power functions were derived separately for force and duration. For force, the exponent was 1.4. However, they noted that because of the tendency for subjects to attenuate the range of numbers used in magnitude estimation (i.e., regression bias; see S. S. Stevens and Greenbaum, 1966), the value of the exponent was depressed and actually closer to 1.7. The latter is also the exponent found experimentally in the earlier study by J. C. Stevens and Mack (1959). Perceived effort also increased with duration, having an exponent of 0.57 (i.e., 0.70 when adjusted for regression bias). One equation could then describe the combination of force and duration which "produces the same terminal perceived force,  $\Psi_t$ :  $\Psi_t = K\Phi^{1.7}t^{0.57}$  where  $\Psi$  represents force in newtons, this time in seconds, and  $K$  is the constant level of perceived force at the limit of endurance. Rearrangement of this equation yields a simplified, one-exponent equation:  $\Psi_t = K\Phi^{2.04}$  which is independent of regression bias, being obtainable from either the biased or unbiased independent estimates for force and duration. Based on these

findings, they suggested that subjects instructed to maintain a constant perceived effort would demonstrate diminishing force exerted over time, a result obtained earlier by Eason (1959).

In a subsequent study, Cain and J. C. Stevens (1971) explored these implications using a unique constant effort task. First, for a group of 12 male subjects, the earlier magnitude estimation study was replicated for phasic squeezes at seven forces between 88 and 429 newtons (19.8 to 96.4 lbs). The obtained power function had an exponent within the expected range. Next, for each of the forces, the subjects were asked to maintain their effort on the handgrip task constant for 2.5 min. Results indicated that actual force exerted typically dropped off rapidly for the first 20 secs, with a much slower decline in force thereafter. A single exception to this generality was for the lowest force squeeze, 88 newtons (19.8 lbs) for which a consistent, slow decline was observed throughout the 2.5 min duration. A two-exponent equation could describe the initially rapid, then slower decline in force across time. On the basis of these results, it was predicted that "two different constant-effort contractions may arouse the same degree of fatigue when the areas under their respective force versus duration curves are equal." A subsequent study impressively supported this prediction.

Since two components were involved in the force-duration curves of Cain and J. C. Stevens (1971) for the constant-effort contractions, a fast then slow decline, it was suggested that two separate physiological processes may be involved in the perception of effort: "(the fast) one—representing activity of mechanoreceptors in the tendons or the skin, and the (slow) term the activity of receptors sensitive to noxious metabolites produced by muscular activity." At least one alternative explanation may be proposed: the fast decline in force may have been due to derecruitment of glycolytic, fast twitch fibers in favor of less rapidly fatiguing oxidative, slow twitch fibers.

Their subsequent studies of surface electromyography (Cain and J. C. Stevens, 1973) indicated the EMG activity in the forearm initially did decrease rapidly similar to force in constant-effort handgrip tasks but tended to asymptote at a point where force was still declining. Their two-process model could also be explained by the alternative explanation noted above. In summary, in a series of elegant psychophysical studies on constant-effort tasks, J. C. Stevens and Cain attempted to relate perceived effort to physiological processes and acknowledged that perceived effort is founded on more than a single underlying physiological process.

In the psychophysical studies described above, perception of force exerted or "Anstrengung" (exertion) was related to stimulus intensity for fixed, brief work periods. Bujas, Paulina, Siemer, Vidacek, and Vodanovic (1966) noted that the research of J. C. Stevens and Mack (1959) and Borg (1962), and presumably Schmale, et al. (1963) and Bernyer (1967), do not

present psychophysical power functions relating the perception of "fatigue" to duration of effort. They pointed out that brief effort or momentary exertion, used in the above studies, may not lead to "fatigue" although the effort expended may be relatively high. Bujas, et al. present information about the psychophysical relationship for the feeling of fatigue *per se* during periods differing in the duration of static work. Using well-trained subjects a steady tension was exerted by holding a 7.17 kg weight with the arm upright and the elbow supported. Durations of 30, 60, 90, 120, 150, and 180 sec. were used and the psychophysical function relating the perception of fatigue to duration of effort was derived using the magnitude estimation method. Fatigue ratings were obtained 30 and 60 sec. after termination of each trial. The subjects were told to include feelings of "tightness, numbness, increase in muscle sensitivity through pain, weight" and so on in their judgments of "fatigue." The obtained power function for the overall group was  $\psi = .0048 (\phi - 30'')^{1.54}$  for the mid-range of the curve relating perception of fatigue to task duration. In a subsequent experiment, the identical procedure was repeated while integrated EMG amplitude from four muscle groups of the arm was recorded. A power function for the mid-range of the curve relating integrated EMG levels to task duration was almost identical to the psychophysical function:  $\Sigma \text{EMG} = .0051 (\phi - 30'')^{1.51}$  indicating impressively high correspondence between the growth of fatigue in static work and EMG activity. It is also of interest that the exponent of the power function for both fatigue and EMG activity in this study falls within the range of exponents observed for force exerted in the studies described above.

These psychophysical studies provide important information about the relationships between perception of effort and load levels during various types of work performance. In general, for positive (i.e. concentric) work, the exponents of the power functions for force exerted and fatigue in a variety of situations fall within a limited range (1.4 to 1.9), indicating that the perceptions increase at a moderate rate relative to changes in the physical work stimulus. An interesting contrast existed for the exponent (3.1) of the weight-holding task of Schmale, et al. (1963) which may be related to the negative (eccentric) nature of the task. However, psychophysical techniques are currently limited in regard to application for the evaluation of subjective fatigue on a longitudinal basis. With the exception of Bujas, et al. (1966), most of these studies relate a subjective quality (e.g. perception of intensity) to the actual intensity of an external physical stimulus. Subjective fatigue can be conceived as arising from a different source; specifically, the cues defining subjective fatigue may derive from the internal physiological processes activated by work. There is thus no easy way to identify intensity of the "physical" stimulus when in fact

precise origins of these physiological cues are not well understood. Presently, there is no clear way to derive the power function describing such relationships. Furthermore, none of the available psychophysical techniques are readily applicable to measurement of subjective qualities during work performance on an ongoing basis. Future applications of psychophysical methods may prove to overcome these difficulties, but changes in methodology will be required.

Alternative scaling techniques discussed in the following sections have been applied more frequently to the measurement of subjective qualities or symptoms during work primarily because of their applicability to prolonged work situations. These techniques may be roughly grouped into (1) nondimensional, single point measures, (2) unidimensional rating scale techniques, and (3) multidimensional rating scale techniques.

#### Nondimensional, Single Point Measures

In Table 11-11, studies using nondimensional, single point measurement of subjective qualities in work are shown. These studies have employed the simplest methods of measurement which generally require a single verbal report during work performance. Most often an attempt is made to identify factors displacing the "fatigue" point toward or away from the initiation of work. Inspection of the table reveals little consistency between the studies in regard to the subjective qualities measured (tiredness and undifferentiated fatigue), tasks used, or subject populations selected. It is worth noting that just such inconsistency has hampered the development of a meaningful overview of research concerning subjective fatigue.

In a study by McGrath, Wittkower, and Cleghorn (1954), the undifferentiated (i.e. global) feeling of fatigue was regarded as an all-or-none event, characteristic of single point measurement. Specifically, ratings of various factors which were regarded as "causing fatigue" during long distance flights were obtained from aircraft crews. The crews studied consisted of Canadian pilots, co-pilots, navigators, and radio officers who flew DC 6's in the Tokyo Airlift. The three legs of the flight were 10 to 18 hours each separated by 48-hour layovers. Group discussions and individual interviews were used to identify factors which the crewmen believed to cause fatigue. The specific factors chosen for inclusion in a questionnaire were identified during these preliminary group discussions and interviews. Each of the thirty-four factors were arranged on a 4-point ordinal scale to permit rating from "no importance" (1) in causing fatigue to "great importance" (4). However, the quantitative data available from the questionnaire was not summarized in the report. Conclusions were based principally on the nonquantitative information derived from the interviews and group discussions.

TABLE 14-II  
NONDIMENSIONAL MEASUREMENT OF SUBJECTIVE FATIGUE

Subjective Quality Measured	Task	Subjects	Studies
"Undifferentiated fatigue"	Airplane flights	Male aircraft crews (n = 100)	McGrath, Wittkover, and Cleghorn (1954)
"Undifferentiated fatigue"	Reaction time task	Male medical students (n = 26) Female P.E. students (n = 15)	Pierson (1963), Pierson and Lockhart (1964)
Tiredness	Manual and office work	Male manual and office workers (n = 376)	Griffith, Kerr, Mayo, and Topal (1950)
Tiredness	Isometric weight lifting	Female psychiatric patients (n = 42)	Hemphill, Hall, and Crookes (1952)

Members of the crews tended to report boredom, but only minimal fatigue, during the first six hours of a flight. From six to 10 hours, the crewmen reported becoming tired, sleepy, and irritable. Causes of fatigue were grouped into three categories: (1) *General Factors*, such as length of flight, delayed flights and false starts, details prior to takeoff, reliability of radio communications and navigational aids, bad weather, monotony and boredom of familiar routes, number of intermediate stops, and drinking the night before the flight; (2) *Specific Factors*, associated with problems peculiar to the DC 6 aircraft (including noise and vibration, design of the flight deck and instrument panels, and uncomfortably fitting oxygen masks), and problems peculiar to the route (including post-flight conditions, recreational facilities at stopover points, and irregular hours); and (3) *Personal Factors*, such as inexperience and tension within the crew, relative burden of responsibility, relationships with higher authorities, and domestic worries.

In this study, no information was obtained to quantify fatigue or performance during the course of flights. The authors were well aware of this basic limitation in their study. They note that the "method used in this investigation might be criticized on the grounds that information gathered was of a purely subjective nature and was given in retrospect." Continuing to discuss these limitations, they clearly indicate the two ingredients minimally necessary to conduct research evaluating the role of subjective fatigue longitudinally during work performance: (1) An adequate *measure of subjective fatigue* for which "no satisfactory objective tests [had] as yet been designed" and (2) *Performance measures* to evaluate "the subject's performance under normal operating [or working] conditions."

Pierson (1963a) attempted to relate the occurrence of "fatigue" in twenty-six male medical students to performance on a reaction time task using an apparatus described by Pierson and Rasch (1959). On each trial, an auditory preparatory signal was given approximately one second before presentation of a light stimulus which cued the response. Each subject was required to release a telegraph key and move their hand 11 in. forward through a light beam when the light stimulus appeared. Reaction time (RT) was measured as the latency between presentation of the light and the release of the key, while movement time (MT) was measured from the release of the key to the breaking of the light beam. Trials were presented at intervals of approximately 10 seconds with the task continuing until the subject indicated that he could no longer proceed. During the task, the subject was required to indicate when he "believed his responses were becoming slower." The trial on which this report was given was defined as the "fatigue" point. Subjects were also asked to indicate the point at which they became bored with the task.



A significant relationship ( $r = .17$ ) was found between RT latency and the trial on which the "fatigue" report was given. This relationship means that slower RT's tended to be associated with later fatigue points, while early fatigue points were more often concomitant with faster RT's. More importantly, no differences in RT or MT were found between five baseline trials, selected near the beginning of the task, five trials following the fatigue point, and the last five trials preceding voluntary termination. Furthermore, no relationship ( $r = .02$ ) was found between the "fatigue" trial and total trials performed before voluntary termination. Pierson (1963b) reported that measures of unspecified isometric strength for these same subjects were found to be unrelated either to RT and MT decrement or the trial defined as the fatigue point. He concluded that "subjective impressions of performance bear little relationship to actual performance" and that the "subjective experience of fatigue is not a valid criterion for the ability to perform speed or endurance type muscular work."

In regard to Pierson's conclusions, it should be noted first that a study subsequently reported by Pierson and Rich (1967) found a steady  $O_2$  consumption during this task. Performance of the task required only slightly greater  $O_2$  uptake than baseline measures obtained while the subjects were sitting quietly (330 ml/min. vs. 280 ml/min.). Therefore, generalization of these results to "speed or endurance type muscular work" appears unjustified. At best, the results may be generalized to sedentary type tasks. Additionally, RT and MT performance decrement was measured as the slowest block of five trials within a sequence averaging 174 trials for the group of subjects. In tasks of this kind, random variations in performance across trials would be expected due to momentary fluctuations in motivation, concentration, and so on. As the measure of performance decrement, selection of the slowest block of five trials may therefore reflect nothing more than random changes in performance. In fact, since there was no systematic trend toward deteriorating performance across trials, as the data indicates, it is hardly conceivable that "the subjective experience of fatigue" would be related to RT or MT decrements in this study. In a word, Pierson's conclusions seem unfounded.

A subsequent study (Pierson and Lockhart, 1964) did qualify the earlier conclusions of Pierson (1963a). Using fifteen female physical education majors with a mean age of 19.6 years, an attempt to replicate the earlier study was made. This time, the principal finding was a *significant* relationship ( $r = .59$ ) between the trial on which "fatigue" was reported and the total number of trials performed. The "fatigue" point for male medical students had been found to be clearly unrelated to endurance ( $r = .02$ ). One might suggest that these conflicting findings may be explained by differences in motivation and cooperation with task requirements. For example, those

medical students who reported "boredom" (7 of 26) did so at a mean trial of 70.4 while those female subjects reporting boredom (4 of 15) did not do so until trial 102.4 even though females quit earlier (an average of 171 total trials for the males vs. 151 for the females). The male subjects evidently quit performing for reasons unrelated to "fatigue," while the female subjects seemed to be more highly motivated, quitting for reasons more related to "fatigue." All other findings in this study were generally consistent with Pierson (1963a). Obviously, there are problems related to replicability as the contrasting principal findings between the studies indicate, and specifically require qualification of the earlier conclusions concerning the relationship between subjective report of fatigue and performance factors.

A study by Griffith, Kerr, Mayo, and Topal (1950) is another example of a single point measurement technique. They employed a simple "tear ballot" used previously by Kerr (1943) to determine the hour during each half of an eight-hour shift when manual workers, foremen, and office workers reported feeling "most rested" or "most tired." To indicate the hour of each half-shift, the subjects simply tore the appropriate corner off the ballot. Data for the study were reduced to percentages reporting "most tired" and "most rested" during each four-hour period. Findings were remarkably uniform for all types of work and are schematized in Figure 11-2. There was a tendency for workers to report feeling "most tired" during the first and fourth hour of the morning and the last hour of the afternoon. Manual workers more than 36 years old reported feeling "most tired" more often during the last hour of the morning and afternoon periods than did younger manual workers. Information about the consistency of this technique is sketchy, although it is reported that "repeat-test reliability coefficients for small groups ranged from .69 to .92." No information was available concerning validity of the self-report since no other measures of work performance were obtained. Information provided by this study is limited, since only a single point of "most tired" was reported, it was impossible to evaluate relative *levels* of fatigue on an hourly basis across the work shifts. By focusing on *percentages* of workers reporting "most tired," an erroneous conclusion was also drawn, i.e. "Older workers report significantly greater variation of such feelings ('most tired' and 'most rested') than do employees under the age of 36." This conclusion may be correct for workers as a whole, but since the relative *degree* of tiredness was not measured, no conclusion can be made concerning the variability in tiredness reported by workers in different age groups.

Hemphill, Hall, and Crookes (1952) also used a single point measurement technique within a simple weight holding task. Using the middle finger of the right hand, female psychiatric patients classified as endoge-

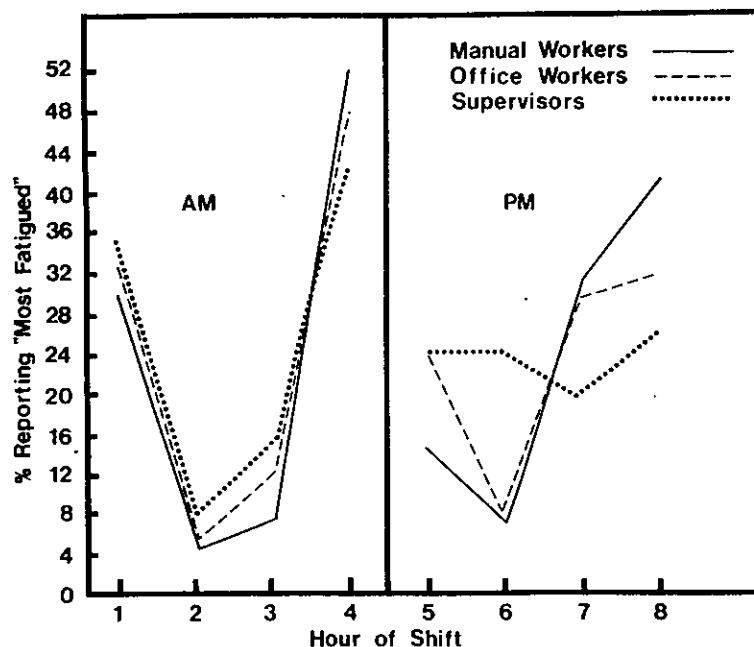


Figure 11-2. Percentage of workers reporting feeling "most fatigued" during the morning (AM) and afternoon (PM) halves of an eight-hour shift. (From J. W. Griffith, et al. "Changes in Subjective Fatigue and Readiness for Work During the Eight-Hour Shift." *Journal of Applied Psychology*, 31:163, 1950).

nous depressives and anxiety neurotics were instructed to lift a small weight (2 kg) for as long as possible. They were asked to report when they "began to feel tired." The results indicated that the "endogenous depressive" patients reported feeling tired more rapidly and quit sooner after reporting tiredness than patients classified as anxiety neurotics. The results suggest that endogenous depressive patients, who are characteristically "fatigued" and with minimal energy reserves, became tired more readily during static work. Anxiety neurotics, on the other hand, are frequently characterized by an agitated, aroused kind of depression which may actually serve to reduce the degree of tiredness during activity or static work.

Single point, nondimensional measures of fatigue are open to criticism. "Fatigue" is conceptualized in these studies as an all-or-none event: either the individual is experiencing a certain "degree" of fatigue, or he is not.

It is hard to defend such a notion, either conceptually or psychometrically. First, experience suggests that, in general, one can report a variety of intensities of subjective qualities—one has, at different times, felt fresh, somewhat tired or extremely worn out. It would seem more logical that the symptoms would increase quantitatively, in some fashion, the longer one performs work. Measurement should be designed to quantify at least ordinal levels of the subjective experience.

Second, from the standpoint of measurement, good psychometrics also suggests that some arrangement be made to measure quantitative *levels* of fatigue. Without this possibility, the relationship between subjective fatigue and work decrement or endurance cannot be clearly specified. It takes at least three points to define a function, and theoretically a subjective event varying in intensity such as subjective fatigue, can be measured on more than three points. Additionally, reliability is known to increase, up to some limit, as the range of scores increases. Single point measures can thus be expected to be less reliable than multi-point, or dimensional, measures of subjective fatigue.

#### Unidimensional Rating Scales

Studies using unidimensional rating scales to evaluate subjective qualities during work are summarized in Table 11-III. An example of an early study attempting to dimensionally quantify a single subjective quality, i.e. "feeling tone," is provided by Foltz, Jung, and Gisler (1944). They had seven male subjects ride a bicycle ergometer at 1235 kpm/min. while pedalling at 54 rpm's until unable to maintain the required level of performance. This work of about 2.8 L. min.  $\text{Vo}_2$  would be approximately 85 percent max  $\text{Vo}_2$  of nearly 3.3 L. min. for an "average" young male. All subjects rode twice daily with the rides separated by a 10-minute rest period until a total of 378 rides were obtained over 27 days. Ratings of "feeling tone" were obtained *before* each ride using a 15-point rating scale ranging from "the worse I have ever felt in my life" (1) to "the best I have ever felt in my life" (15) while the midpoint (8) required the response "the way I usually feel." Ride duration was the measure of performance and ranged from 2 to 22 min. The study was designed to test the hypothesis that work performance is influenced by how an individual feels before beginning work. Results indicated that the relationship between ride duration and prework "feeling tone" was negligible and, in fact, slightly opposite from predicted. The authors suggested that work performance was unrelated to prework subjective feeling states.

The approach used by Foltz, et al. represents one of the early systematic efforts to relate a subjective quality (prework "feeling tone") to performance in a laboratory work situation. However, this study again exemplifies

TABLE 11-III  
RATING SCALES IN THE MEASUREMENT OF SUBJECTIVE FATIGUE

Subjective Quality Measured	Scale Point Range	Task	Subjects	Studies
"Feeling tone"	13	Bicycle ergometer	Healthy males	Foltz, Jung, and Ciesler (1944)
"Undifferentiated fatigue"	4	Housework	Healthy females; housewives	Gross and Bartley (1951)
"Undifferentiated fatigue"	9	Bicycle ergometer; normal conditions	Healthy males	Nunney (1963)
Tiredness	5	Mental work, Verbal and qualitative	Healthy males	Poffenberger (1928)
Tiredness	20	Psychomotor tasks; drug states	Healthy males; US Air Force enlisted men	Pearson and Byars (1956); Pearson (1957)
Tiredness	20	Prolonged aircraft flights	Healthy males; aircraft crews	Buckley and Hartman (1969); Hartman, et al. (1973a, b)
Tiredness	9	Psychomotor task	Males and females; college students	McNelly (1957)
Tiredness	5	Factory workers	Healthy males; butchers	Bufas, Strenge, and Vidacek (1965)
Tiredness	10	Portages	Healthy males	Strauss and Carlock (1966)
Tiredness	7	Bicycle ergometer	Healthy males	Heuting and Sarphati (1966)
Tiredness	13.5	Psychomotor task; visual threshold	Healthy males; college students	Walster and Aronson (1967)
Perceived effort	15	Bicycle ergometer; normal conditions	Healthy males	Boyg (1961a); Boyg (1962); Boyg and Linderholm (1967); Frankenhauer, Post, Nordhafer, and Sjolberg (1969); Kay, and Shephard (1969); Bar-On, et al. (1972); Skinner, et al. (1968; 1973b); Morgan (1973)
Perceived effort	15	Bicycle ergometer; normal conditions	Male cardiac patients	Boyg and Linderholm (1970)
Perceived effort	15	Bicycle ergometer; hypoxic conditions	Healthy males	Gerbien, House, and Wilsman (1972); Weiser, et al. (unpublished observations)
Perceived effort	15	Bicycle ergometer; dynamic weight lifting; wheelbarrow pushing	Healthy males	Gamberale (1972)
Perceived effort	15	Bicycle ergometer; concentric and eccentric work	Healthy males and females	Honrikson, Knutigen, and Bonde-Petersen (1972)
Perceived effort	15	Treadmill walking and tanning	Healthy males	Noble, et al. (1973)
Perceived effort expended	5	Treadmill walking	Healthy males	Lloyd and McCluskey (1971)
Perceived effort expended	5 and 10	Isometric contraction; hand dynamometer	Healthy males	Caldwell and Smith (1967); Caldwell (1967)
Pain intensity	5	Isometric contraction; hand dynamometer and arm pull	Healthy males	Caldwell and Smith (1966); Caldwell (1967); Meurer, Smith, and Caldwell (1968); Lloyd, Voor, and Thrieman (1970); Lloyd (1971); Lloyd (1972)
Discomfort	4	Bicycle ergometer	Healthy males	Gagge, Stoltwijk, and Salin (1968)
Thermal sensation	7			

the caution required in design and interpretation of research in this area. First, no attempt was made to equate work levels for the individual subjects. While information was not presented regarding individual differences in such factors as max  $\dot{V}O_2$  or heart rate response to specific work levels, almost surely there were differences in working capacity between the individual subjects. These individual differences in working capacity, while not well known at that time (Simonson and Enzer, 1942), would greatly influence ride duration. Failure to equate work levels during exercise, however, would lead to a substantial attenuation of the relationship between "feeling tone" before work and subsequent work performance.

Second, it would appear that performance data for all 378 rides and subjective ratings for the seven subjects were grouped across subjects to obtain the correlation between ride duration and prework "feeling tone." Grouping data in this way violates an assumption of independence of the pairs of observations for correlational analysis and could further attenuate any potential relationship in the data. Therefore, on the basis of this study, little can be concluded regarding the relationship between pre-work "feeling tone" and subsequent work output due to these inherent design and methodological problems. In addition, since subjective symptomatology *during* work, or produced by work, was not measured, the implications for the area of subjective fatigue are limited. However, studies of this sort could lead the unsuspecting reviewer to conclude that "research has indicated" that subjective symptoms are unrelated to work performance.

In contrast to Foltz, et al., the remaining studies described in this section measure subjective qualities *during* work performance. These have focused upon five types of subjective qualities during work: undifferentiated fatigue, tiredness, effort, pain, and thermal discomfort.

#### *Undifferentiated Fatigue*

A study by Gross and Bartley (1951) provides an example of the use of a simple rating scale to measure undifferentiated fatigue. A four-point rating scale (Great, Moderate, Little, None) was employed. A group of twenty women rated their level of fatigue at 30 minute intervals at the beginning and throughout a two-hour period of housework. Records were maintained concerning the quality and quantity of work performed. Those reporting the most fatigue cleaned fewer rooms than those moderately fatigued, but more than those reporting little or no fatigue. The women reporting the most fatigue also received the lowest ratings for quality of housework performed. Additionally, it was found that the rating of fatigue for some women tended to peak at the middle of the work period even though there was no intervening rest. The latter observation was regarded as support for the notion that subjective fatigue is not consistently

related to the amount of energy expended. Examination of the individual protocols suggests that some subjects may have been timing the work period, reporting less fatigue as the end of the period neared. In contrast, most reported increased fatigue toward the end of the work period. This is an occasional observation when work periods are of a fixed duration, and appears related to expectations about the length of a dull, nonrewarding task. Walster and Aronson (1967; see p. 365) found an increase in fatigue reports toward completion of a monotonous task of fixed duration. Additionally, Gross and Bartley noted that there was a "lack of distinction in the minds of the subjects between localized muscle discomfort from stooping, etc., and the overall personal experimental (sic) state we identify as fatigue." Recently, a task specific category of subjective qualities experienced during work, including those related to localized muscle discomfort, have been identified statistically within a larger set of fatigue symptoms (Weiser, Kinsman and Stamper, 1973).

Nunney (1963) used a nine-point rating scale to evaluate subjective fatigue in eighty male college students. The testing program encompassed six test sessions of five min. given on alternate days. The subjects were divided into five groups: A control group; two groups who bicycled on an ergometer, either with no load or a seven-lb load; and two groups who ran on a treadmill at 6 mph, either on a 0 percent or on a 25 percent grade. The rating scale ranged from No Fatigue (1) to Extreme Fatigue (9) and was given 15 sec. after cessation of work. Pulse rate was measured before, at the end of the work task, and after 10 min. of recovery. If the mean testing pulse rate was 80 bpm (the actual value was not reported), then the mean pulse rate was 91 and 137 bpm for the 0 and seven-lb bicycle work tasks and 141 and 181 bpm for the 0 percent and 25 percent treadmill work. This indicates that the work loads were progressively heavier, with the two middle tasks nearly the same and the heaviest nearly 100 percent max  $\dot{V}O_2$ . We note that Nunney, similar to Foltz, et al. (1944) did not control for individual differences in the ability to do aerobic work as measured by max  $\dot{V}O_2$ . The fatigue scale scores at the end of work were 1, 4, 4, and 6 for the respective work loads. The *group means* for pulse rate changes and fatigue scale scores correlated highly ( $r = .99$ ), suggesting that for tasks demanding a larger circulatory adaptation there is a corresponding increase in subjective fatigue. Intratask variability was large for fatigue scale scores, but these scores were not significantly correlated to pulse rate changes. Considering the relationships between pulse rate and subjective fatigue ratings for a single work load on a task, the correlations were low ( $r$ 's ranging from  $-.03$  to  $+.37$ ). These low correlations would be expected when the range of work loads relative to max  $\dot{V}O_2$  within tasks are attenuated relative to the range represented across all tasks and work loads. While

Nunney concludes "that physiological changes are not directly related to fatigue," his data actually support a direct relationship between fatigue measures and work as indicated by heart rate changes when there is a sufficiently high range in work loads.

### Tiredness

As early as 1928, Poffenberger reported the use of a simple rating scale to evaluate the feeling of tiredness during prolonged *mental* work. Ten subjects performed up to fourteen consecutive trials on each of four tasks: continuous addition, sentence completion, judging of written compositions and various independent forms of a rather lengthy intelligence test. In the case of the composition judging, 10 specimen compositions were rated on each trial. For the other tasks, each trial required 20 to 30 minutes of continuous performance. Performance was measured on each trial, and ratings of tiredness were obtained using a seven-point *ordinal* scale, given immediately before the first trial and after each trial of a task. The rating scale ranged from "extremely good" (1) to "extremely tired" (7) with a mid-point defined as "medium" (4). Results indicated that the mean rating of tiredness increased across trials in a fairly linear manner for all tasks, although overall performance remained (1) unchanged—for sentence completion or composition judging, (2) improved—for intelligence testing, or (3) deteriorated—for addition. Due to improvement in the tasks resulting from practice, the relationship between mental work and the feeling of tiredness may have been masked for all of the tasks except addition. Presumably, addition is a highly learned task, the performance of which would not be expected to improve with additional practice. Thus, the relationship between the subjective quality "tiredness" and performance would be most clearly manifest for the addition task. Subjects showing the greatest performance decrement for each task reported a large increase in tiredness; those showing the least performance decrement reported substantially less tiredness across trials. Poffenberger's results thus indicate a fairly clear relationship between "tiredness" and performance in prolonged mental work.

Pearson and Byars (1956) and Pearson (1957) reported the development of a 10-item unidimensional scale measuring the subjective quality of tiredness. Their *interval* scale represents the most thorough effort to devise a measure of subjective fatigue as a unidimensional feeling of tiredness. Procedures for scale construction recommended by Edwards and Kilpatrick (1948) were followed. The problem was to select a set of adjectives that could be arranged in a continuum, with *equidistant* adjacent items, so that the subject could report his level of tiredness by pinpointing a position along this continuum. An initial set of 500 adjectives were selected. All am-

biguous items were eliminated, and the remaining ninety-two items were scaled using Thurstone's method of equal appearing intervals (see Guilford, 1954). Briefly, independent judges sorted each item into nine categories ranging from extreme well being (1) to extreme fatigue (9). All forty-eight items showing low interjudge agreement were eliminated. The forty-four remaining items were arranged in a checklist and provided with choices of "better than," "same as," or "worse than" allowing an individual to place himself on a continuum of subjective tiredness. A developmental study involving a lengthy psychomotor task (Multidimensional Pursuit Test; Haury and Payne, 1956) was used to further eliminate those items with low validity and internal consistency. Finally, two equivalent forms (A and B) of the checklist were constructed. The item composition of one form (Form A) of the checklist is shown below:

Item	Order No.
1. Like I'm bursting with energy	(2)
2. Extremely peppy	(6)
3. Very lively	(12)
4. Very refreshed	(9)
5. Quite fresh	(4)
6. Somewhat fresh	(7)
7. Slightly tired	(1)
8. Slightly pooped	(5)
9. Fairly well pooped	(11)
10. Petered out	(8)
11. Very tired	(13)
12. Extremely tired	(3)
13. Ready to drop	(10)

In application, the items were arranged in a random order as indicated in the right-hand column. The colloquial content of the checklist ("petered out," "ready to drop," "fairly well pooped") is apparent.

A validation study using the same psychomotor task showed that equivalent forms of the checklist intercorrelated highly ( $r = .92$ ) both for 100 experimentally fatigued and for 100 control subjects. The scores for experimentally fatigued subjects were significantly increased from pre- to post-task. Additionally, the checklist clearly differentiated between experimentally fatigued and control subjects. Subsequently, the checklist was shown to differentiate between subjects given analeptic (5 mg of dextro-amphetamine sulfate) and depressant (.65 mg hyoscine hydrobromide mixed with 50 mg of diphenhydramine hydrochloride) drugs in a single-blind study. As a unidimensional measure of tiredness, Pearson's interval scale is unsurpassed since it was constructed systematically according to sound psychometric principles.

Unfortunately, application of the checklist beyond the original validating studies has been limited although the authors suggest that its applicability should be general. Its widest application has been to evaluate changes

in fatigue during aeromedical studies by Hartman and his coworkers. In a study by Buckley and Hartman (1969), Pearson's checklist was used to evaluate changes in subjective tiredness during a 32-hour transatlantic helicopter flight. The flight was accompanied by loss of usual sleeping patterns, decreased food intake, continuous vibration, and flicker illumination common to rotary wing aircraft. Ratings on tiredness showed an almost linear increase throughout the flight, with an interesting recovery spurt during fly-by in a Paris airshow even though no rest periods intervened. During this fly-by, which occurred at the end of the flight, tiredness ratings decreased to a point reported 14 to 16 hours earlier. Such a recovery phenomenon may occur in tasks of fixed duration in which there is high expectation of personal reward upon successful completion. Such a phenomenon may be related to the Hullian goal-gradient (Hull, 1913; 1952). Similar application of this scale have been made to the study of the in-flight fatigue of FB-111 crews (Hartman, Hale and Johnson, 1974) and C-5 jet transport crews (Hartman, Hale, Harris, and Stanford, 1974).

McNelly (1954) also used Thurstone scaling techniques to develop a 9-point rating scale of tiredness. Again, an attempt was made to construct an interval scale with adjacent scale points of equidistance. One form of McNelly's scale is shown below:

Item	Mean Scale Point
About to fall over	1.1
Lagged	2.2
Let down	3.2
A little tired	3.8
Average	5.0
Fairly well	5.9
In gear	7.0
Very good	7.9
Terrific	8.9

Preliminary Thurstone scaling was performed on 123 adjective items or phrases by 10 judges. The judges sorted each item into categories according to increasing degree of tiredness, establishing a preliminary mean scale point value for each item. Sixty (60) adjective items or phrases with the least variability were retained, and rescaled by the same judges. The second scaling had a high correlation ( $r = .95$ ) with the first scaling. Of these 60 items, two groups of nine items each were selected on the basis of (1) having low variability and (2) being approximately one scale point apart. The mean scale values for one set of nine items are shown above.

The validity of the scale was tested on eighty male and female undergraduates using a block turning task with heavy (11 ounces) or light (.4 ounces) blocks either with massed practice (without rest) or spaced practice (rest after each trial). Each trial consisted of turning twenty blocks as rapidly as possible. The interval scale was administered before and after

1,000 trials. As expected, spaced practice produced both less work decrement and lower "tiredness" scores than did massed practice. Ratings of tiredness showed a significant change pre- to post-work, although the relationship between the subjective ratings and work decrement was quite low ( $r = .30$ ). While McNelly's scale has frequently been cited, to the best of our knowledge it has not been used in any other studies reported in the literature. Its validity in other work situations is essentially unknown. Similar to the Pearson and Byars' scale (1956; Pearson, 1957) the composition of McNelly's scale includes colloquial items (e.g. fagged, in gear, etc.) which appear to be unique in regard to culture, time, and geographic location.

Bujas, Sremeg, and Vidachek (1965) comprehensively investigated the relationships between the subjective quality of tiredness and other variables in 115 butchers from the same factory after an eight-hour work shift. A five-point scale was used to obtain ratings of tiredness. The scale ranged from "Generally I am not tired" (1) to "I feel very tired" (5). Only 4.4 percent of the workers reported that they were generally *not* tired after work, while 26.5, 49.6, 12.4, and 7.1 percent reported feeling "slightly," "moderately," "quite," or "very" tired, respectively. Ratings of tiredness were independent of estimates of the workers' productivity made by the foremen and plant supervisors. However, the authors note that productivity ratings tended to be uniformly high for all workers so that no discrimination in productivity was achieved by the supervisors' ratings. It is interesting to note that the tired workers reported having less satisfaction with their work, sleeping less well at night, less appetite, generally feeling less healthy, and more often being in a worse mood at and away from work than their less tired co-workers. This study is similar in many ways to that of McGrath, et al. (1951) in that the emphasis was upon identification of factors *contributing* to the subjective feelings of tiredness or fatigue that occur during occupational activities.

Strauss and Carlock (1966) measured subjective tiredness during portages with various loads using an 11-point rating scale developed by Psychometric Associates (1954). This rating scale ranged from "worn out, too tired to do anything" (1) to "not tired at all, fresh enough to start a full day" (11). Ten male subjects ranging in age from 18 to 35 years carried weights of 10 to 34 lbs during three walks over a two-mile course performed at two-hour intervals. Two portage modalities were used for each weight condition: comfortable, i.e. weights carried on the back, and uncomfortable, i.e. weights carried by hand. Performance on a battery of perceptual and psychomotor tests and ratings of tiredness were obtained before and after each portage. Walking time did not differ between light and heavy portages and was found to be unrelated to tiredness scores. However, the most uncomfortable portage modality was associated with reports of more tired-

ness. Task performance was significantly related to ratings of tiredness for each of the tests administered, with correlations ranging from  $-.62$  (for perceptual speed and accuracy) to  $-.89$  (for locating and marking identical pairs of names within a list).

Heuting and Sarphati (1966) noted that subjective ratings obtained during work performance could be subject to Titchener's "stimulus error." Specifically, the load levels on an ergometer may be presented in such a way as to provide the subjects with *external* information capable of altering their judgments about subjective tiredness. Thus, a subject given information about the load level at which he is performing will be apt to base his subjective rating of tiredness, in part, on that external information. Heuting and Sarphati had their subjects perform on a bicycle ergometer for 11 minutes on 13 successive days. The terminal load level for each daily session was selected as one of seven equal increments between 1050 and 1950 kpm/min. For each session, the initial work load was 45 kpm/min; it was increased by 10 percent of the final load level for each minute thereafter. Previous studies (Heuting, 1964; Heuting and Visser, 1960) indicated that the work load appeared to increase equally for each of these graded work schedules. Thus, estimates of tiredness could be obtained that were less affected by external information. Tiredness was rated in three ways: (1) during each minute of work the subjects positioned a potentiometer pointer along a blank scale to a location between horizontal and vertical corresponding to their level of tiredness; (2) immediately post-work the subjects adjusted the volume of a white noise source to match their level of tiredness; and (3) immediately post-work, the subjects indicated their level of tiredness on a seven-point scale ranging from "feeling fit, rested" (1) to "feeling extremely tired, exhausted" (7).

Average correlations between the terminal work load and post-work tiredness ratings ranged from  $.71$ , for the pointer, to  $.49$  for both the white noise and seven-point rating scale. *Individual* correlations generally ranged from  $.54$  to  $.93$ . The regression lines between work load and individual post-work ratings of tiredness were clearly linear. The study indicates the feasibility of measuring subjective status during work performance by rating scale techniques. Additionally, the potentiometer method showed the highest post-work relationship to work level. This result may have been due to the subjects' familiarity with this rating technique since it was used more frequently than the other two methods. Thus, there is the suggestion that *experience* with the rating technique may improve the relationship between subjective status and work load.

Janssen and Docter (1973) explored the change in fatigue during static work and recovery using Heuting's method (1968) of "fatigue" assessment. They had six well-trained male subjects exert pulls with the forearm at

25, 35, and 45 percent of their voluntary maximal capacity for intervals of 2.5 min., followed by 2 min. of recovery. Fatigue measures were obtained at rest and at 30 sec. intervals throughout the static work task and recovery. An essentially linear growth of fatigue was observed, more rapid for the higher work loads, with the expected steeper decline during recovery. Reproducibility was very high, and similar results were obtained for the left and right arm. Predictably, heart rate changes in this static work task were less impressively related to the course of work and recovery than the fatigue ratings (see Chap. 11 "Physiology of Work Capacity and Fatigue"). The impressive linearity may have been partially a result of timing during this relatively brief, fixed duration task.

While Heuting and Sarphati (1966) noted that external information about the work load level during exercise could influence reports of subjective tiredness, they did not directly test the effects of such external information on subjective reports. A more direct test of the effects of task information on subjective report was made by Walster and Aronson (1967). They reasoned that reports of tiredness would depend on time perceived to be remaining in the task: after the same duration on a task, a report of "very tired" would be incompatible with information that considerably more time is required. In contrast, such a report would be more compatible with information that the task is near completion. Two groups of 10 subjects were told that they would be required to complete either three (short expectancy) or five (long expectancy) trials during the study. Each trial involved two tasks: (1) marking X's on graph paper squares at the rate of one/second for 10 minutes, and (2) obtaining the visual threshold to white light which required approximately 10 additional minutes. Ratings of tiredness were obtained on each trial after the X-marking task. A rating scale ranging from "as fresh as I have ever been" (1) to "as tired as I have ever been" (13.5) was used. Only data from the first three trials were used in making the comparison between groups. Results indicated that the subjects in the three-trial condition reported a sudden increase in tiredness on trial three, while subjects in the five-trial condition did not report a corresponding increase. These results support the notion that information about task duration can influence subjective report. As yet, there have been no studies testing this effect for tasks involving physical work. This could be done by manipulating external information about such variables as load level and task duration. It would also be of interest to explore the possibility that motivational factors would differentially affect the subjective reports near the end of the task.

Subjective reports of fatigue may well be affected by the relationship between external (e.g. apparent task duration) and internal physiological cues as described by Snyder, Schulz, and Jones (1974). Briefly, their hypothesis

was that when an internal standard of subjective fatigue is lacking, individuals believing that they have worked longer durations demonstrate more behavioral fatigue than those believing they have worked shorter durations, even though both groups work equally long. However, when clear, physiological cues are present, the effect of *apparent* duration may well be reversed since it conflicts with the predominant internal standard. A study using brief psychomotor tasks appeared to support this hypothesis. Studies such as these deriving from attribution theory (see Jones, Kanouse, Kelley, Nisbett, Valins, and Weiner, 1972) could contribute substantively to subjective fatigue and its behavioral correlates. Unfortunately, very little has yet been accomplished in this regard, and more adequate subjective rating techniques need to be employed together with improved physiological measures and work performance tasks.

### Exertion

During work performance, rating scale techniques have been used to evaluate *perceived effort* by Borg (1962) and his co-workers while *perceived effort expended* has been studied by Lloyd and McClaskey (1971). Both will be discussed below.

**EFFORT.** In a recent review, Borg (1973) has traced the development of a rating scale of perceived effort (RPE) specifically designed for use during bicycle ergometer work. Initially (Borg, 1962) the scale consisted of 21 points with odd points, 3 through 19, anchored by verbal phrases: 3, extremely light; 5, very light; 7, light; 9, fairly light; 11, neither light nor laborious; 13, fairly laborious; 15, laborious; 17, very laborious; and 19, extremely laborious. On the basis of several initial studies, the scale was revised to include 15 points, 6 through 20, as follows:

6	
7	Very, very light
8	
9	Very light
10	
11	Fairly light
12	
13	Somewhat hard
14	
15	Hard
16	
17	Very hard
18	
19	Very, very hard
20	

The scale was presented in quarto fashion, with equal intervals between adjacent scale points, so that the full scale could be seen by the subject when making his ratings of perceived effort. The RPE scale is currently "so constructed that the heart rate (HR) of a normal, healthy middle-aged

man can be predicted if the RPE value is multiplied by 10; thus  $RPE \times 10 = HR$ " (Borg, 1971). Borg cautions, however, that this equation should not be taken too literally since deviations from the equation result from "difficulties in rating, lack of motivation, and disease." Specific circumstances in which the relationship between HR and RPE can be altered will be discussed below.

The load level giving a heart rate of 170 beats/min. can be obtained to define the Physical Working Capacity<sub>170</sub> (PWC<sub>170</sub>; Wahlung, 1948). At a work level corresponding to PWC<sub>170</sub>, an RPE of 16.5 was obtained and defined as PWC<sub>R</sub> (Borg, 1962a). The correlation between PWC<sub>170</sub> and PWC<sub>R</sub> was .61 indicating that, in general, the higher the work load for PWC<sub>170</sub>, the higher the PWC<sub>R</sub>. For an overall work test, RPE tended to grow as a negatively accelerated function of the work load. The correlation between HR and RPE across several work loads was .85, indicating exceptionally high validity. The high validity suggests that the reliability of the rating scale would also be high since, as a general rule, validity cannot exceed the square root of reliability (Cronbach, 1960). This is to be expected since the linear work load-heart rate relationship for individuals has been shown to be highly reliable (Rowell, 1969).

In a validation study, PWC<sub>R</sub> and PWC<sub>170</sub> were compared to wage rates among forestry workers (Borg, 1962a). The PWC<sub>170</sub> correlated at a low level ( $r = .24$ ) to wage rates, while PWC<sub>R</sub>, the subjective estimate, was considerably higher ( $r = .54$ ). Borg suggests that the PWC<sub>R</sub> takes into account motivational components not inherent to PWC<sub>170</sub> but which affect on-the-job performance considerably.

Gamberale (1972) has compared RPE on other tasks with that obtained during bicycling. He noted that the RPE scale was constructed to reflect heart rate (HR) increments of 10 beats between successive scale points. According to this basis, a rating of seven on the RPE scale ("Very, very light") would correspond approximately to a heart rate of 70 beats/min. Twelve male subjects (20 to 35 years of age) performed three tasks: (1) bicycling at work loads of 300, 600, and 900 kpm/min; (2) repetitive lifting of 1.35, 3.35, and 5.35 kg weights 25 cm above a shoulder height starting point; and (3) pushing a wheel barrow loaded with 36, 66, and 96 kg at 100 m/min. In general, these tasks have different muscle mass requirements. The bicycle ergometer task may be characterized as dynamic work requiring a large muscle mass, while the weightlifting involved work using a small muscle mass. The wheel barrow task was a combination of static work (load on the arms) and dynamic work (walking).

The range in heart rates and RPE ratings was greatest for the bicycle ergometer task (approximately from 100 to 185 beats/min), and least for the wheel barrow task (from 100 to 130 beats/min). The weight lifting pro-



duced an intermediate heart rate response (from 110 to 140 beats/min). Mean RPE and HR values were linearly related for all tasks with the RPE close to 1/10 of HR. However, RPE correlated most highly with HR for the bicycle ergometer task ( $r = .94$ ), least for the wheel barrow task ( $r = .42$ ), and at an intermediate level for the weight lifting task ( $r = .64$ ). In general, instructions to rate the overall perception of effort, as opposed to rate specifically the perception of efforts in arms or legs, resulted in a higher relationship between the mean HR and RPE. Nevertheless, differences were observed in the relationships between RPE and HR with specific and overall instructions. As might be expected, for the weight lifting task, specific instructions resulted in a closer relationship between RPE and HR. This suggests that caution should be observed when instructing subjects to employ the RPE scale for different tasks.

Gamberale noted that "it proved beyond the scope and the possibility of the present investigation to provide a satisfactory estimate of the level of correlation between heart rate and RPE." On a statistical basis, assuming equivalent variability in the measure of RPE and heart rate, it could be expected that a reduction in range of heart rate, as in the wheel barrow and weight lifting task, would lead to an attenuation of the correlation between heart rate and ratings on the RPE scale. It would be worthwhile to evaluate the relationship between RPE and heart rate on tasks differing in the degree of dynamic work and muscle mass involvement while employing equivalently sizeable ranges in heart rates produced across levels of each task.

As will be discussed (see p. 391), within the past decade a substantial international literature has become available concerning the systematic measurement of a specific subjective quality, perceived effort, in carefully controlled work conditions (also see Table 14-III).

**EFFORT EXPENDED.** Another aspect of subjective effort is the perceived amount of total available effort expended during work performance. In a study by Lloyd and McClaskey (1971), eighteen male subjects with an average age of approximately 21 years rated perceived effort expended during performance on a motor-driven treadmill using a unique five-point scale. The rating technique had been developed earlier by Caldwell (1967) for use on an isometric handgrip task (see p. 370). Each subject was asked to respond when he perceived that successive 1/5's of his total available effort had been expended. The first response during work (1) was defined as the point where 1/5 of his total available effort had been expended; the response (5) defined the point at which he had expended *all* of his effort. Thus, the rating procedure can be described as a self-paced technique in that ratings were volunteered by the subject rather than requested at selected points by the experimenter. By prior agreement, performance terminated when the sub-

ject called out "5". During the task, all subjects walked on the treadmill at 0 percent grade at 75 percent of their maximum walking speed. Maximum walking speed was determined by noting the points at which the subject broke into a jog during three preliminary trials with increasing treadmill speed and the point at which he started walking after jogging during subsequent deceleration. Eight daily walks at 75 percent maximum walking speed were obtained for each subject during which he called out the points indicating when successive 1/5's of his total available effort had been expended. All subjects were instructed to continue walking as long as possible.

As shown in Figure 14-3, results indicated a highly linear relationship between subjective judgments of the amount of effort expended and mean time on the task. Endurance times tended to increase through the third session, and showed a slight decrease on subsequent sessions. This increase in walk duration may have been due to the effects of practice. The degree of linearity between subjective ratings and mean task time generally increased across the eight sessions, with the last five sessions showing remarkable linearity. In view of these results, Lloyd and McClaskey suggest that subjects "have both the capacity to attend to psychological and physiological cues and to report, on a psychological continuum, their judgments of increasing difficulty of the task or their diminishing capacity to perform."

While the degree of linearity in the relationship between perceived ef-

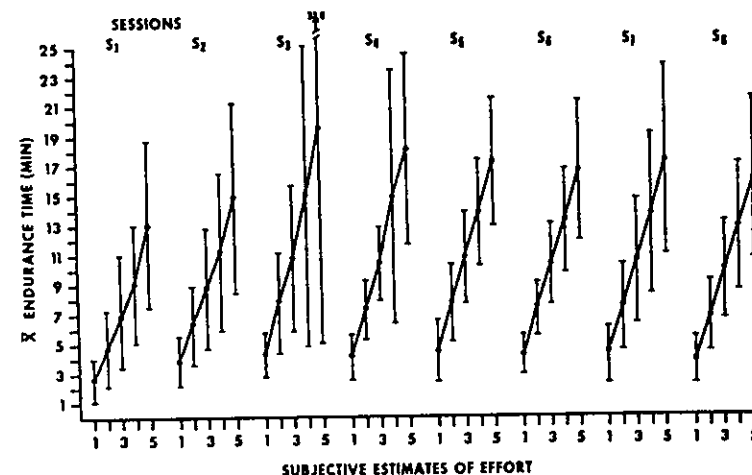


Figure 14-3. Subjective estimates of effort related to mean endurance time during treadmill performance (From A. J. Lloyd, and E. B. McClaskey, "Subjective Assessment of Effort in Dynamic Work," *Journal of Motor Behavior*, 3:49, 1971).

fort expended and task time is impressive, the study raises several interesting questions. First, considerable variability was represented in the data. Standard deviations shown in Figure 14-3 indicate a very broad range of endurance times. In addition, average endurance times (approximately 13 to 19 min.) seem short for such a treadmill task. Were the subjects really performing near the limits of their endurance capacities? In fact, Lloyd and McClaskey note that on session three, which had the greatest variability, two subjects were able to walk 41 and 69 minutes. This raises the possibility that having subjects perform at a given percentage of maximum walking speed might not effectively place each subject at the same relative work load. In other words, 75 percent maximum walking speed may be a physiologically "empty" variable for normalizing endurance capacity compared to maximum aerobic power or other physiological variables such as heart rate. Unfortunately, no data is available in the literature examining the relationship between percent maximum walking speed and measures of endurance capacity. Second, to what degree would the observed linearity found for the subjects as a group exist for individual subjects? No data is presented in this regard. Third, to what degree does this task represent an exercise in timing? The possibility of a timing artifact can be raised almost everytime the self-paced rating technique is used. While the authors clearly regard the data as a demonstration that perceived effort expended is being measured, it is quite possible that each subject was simply *timing* successive fifths of the total duration he was *willing* to spend on the task. Stevens (1957) has discussed categorical and magnitude estimation of time durations which humans can learn to do quite well. No doubt perceived effort expended was being measured in some way, but it is difficult to separate the degree to which effort expended contributed to these results apart from judgments of elapsed time. Additional comment will be offered in the following section regarding timing artifacts in the self-paced rating techniques.

### Pain

Pain is another subjective quality experienced during work performance. Caldwell (1967) and Caldwell and Smith (1967) have found a linear relationship between time on an isometric task and ratings of both pain intensity and perceived effort expended. In both of these studies, the subject was required to hold a hand dynamometer at submaximal levels until it became necessary, in the subject's own judgment, to stop. Subjects were required to indicate when pain intensity attained values of 1 to 5 in the same self-paced fashion as that used by Lloyd and McClaskey (1971) for ratings of perceived effort expended. Under these conditions, ratings of pain intensity could be influenced by the subjects' ability to time successive intervals,

as has been discussed. If, with practice, a subject could call out successive ratings of pain intensity at intervals matched on the time dimension, the relationship between ratings of pain intensity and time on the isometric task would attain a spurious linearity. As noted above, this is a problem also relevant to the interpretation of the Lloyd and McClaskey (1971) study.

Menzer, Smith, and Caldwell (1969) addressed the problem of the possible timing artifact directly. In this study, using the same isometric task, pain ratings by the self-paced technique were compared to those obtained by an irregular report technique adapted from Beecher (1966) and Smith, Egbert, Markowitz, Mosteller, and Beecher (1966). As in previous studies, the subjects were required to sustain a submaximal handgrip, either 25 or 40 percent of maximum, until they judged it necessary to stop. One group of subjects used the self-paced technique calling out the numbers "1" through "5" as these subjective intensities were reached, terminating performance at the highest rating of pain intensity, "5." A second group of subjects were required to select a rating of from "1" to "5" at the request of the experimenter made at random, irregular intervals during the task. All subjects were practiced on *both* rating procedures *before* the experiment began. Their notion was straightforward: if timing of the intervals accounted for the linearity observed in previous studies, the irregular report ratings would be expected to be less linear than the self-paced technique. The results, however, clearly indicated that for subjects practiced on both techniques, the self-paced and irregular ratings produced equally linear relationships with time on the isometric task for average values of perceived pain for the group as a whole.

Nevertheless, these results do not eliminate the possibility that self-paced ratings of perceived *effort expended* used by Lloyd and McClaskey (1971) contained a timing artifact. In the latter study, each rating scale point referred to the estimated proportion of total available effort expended. This may be worthwhile reviewing in some detail. A rating of "1" meant that 1/5 of the subject's total effort was expended; when the subject called out "5" this meant, by definition, that *all* his available effort was expended and he stopped performing the treadmill task. Such a procedure is virtually identical to having the subject state when he is one-fifth through the task, two-fifths through the task, and so on, making the successive ratings highly dependent on the subject's ability to estimate and match successive intervals.

In an excellent study by Lloyd, Voor, and Thieman (1970) ratings of pain were obtained during an isometric hand dynamometer task of either 25 or 50 percent of maximum grip strength. During each of two trials, the forty male subjects reported their ratings of pain intensity using the self-paced technique. The two trials were separated by a rest period of 15 minutes. The results of the study are depicted in Figure 14-4, showing the re-

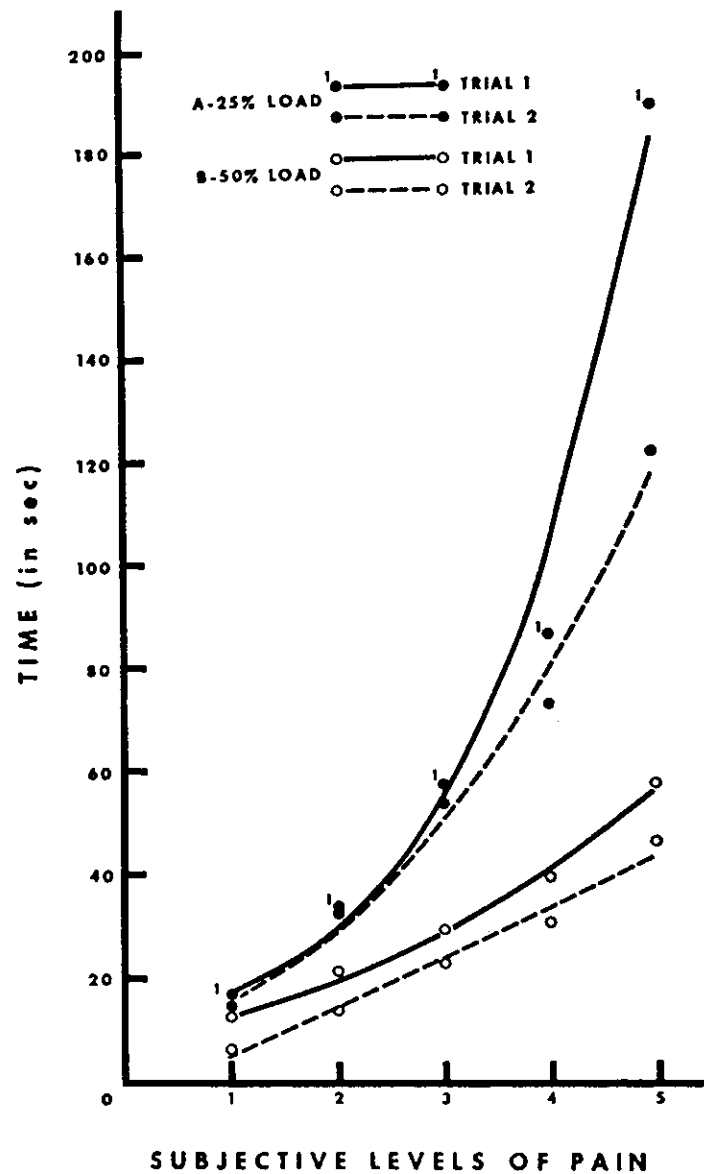


Figure 14-4. Subjective estimates of pain related to endurance time during isometric hand dynamometer tasks (From A. J. Lloyd, J. H. Voor, and T. J. Thiemann, "Subjective and Electromyographic Assessment of Isometric Muscle Contractions," *Ergonomics*, 13: 685, 1970).

relationship between pain intensity ratings and average time on the isometric task. From the figure, it can be seen that ratings of pain increased regularly across time on the task, in a generally quadratic fashion, particularly for the 25 percent load level. Ratings for the 50 percent load level were much more linear for both trials. It is interesting that Trial 2 was more linear in appearance than Trial 1 for both load levels, although the decrease in average task duration for the 25 percent load level on Trial 2 was more pronounced. A decrease in the average task duration for the 25 percent load level may reflect a learned timing effect inherent to the self-paced ratings of pain. This effect due to timing of successive intervals, was not observed by Menzer, Smith, and Caldwell (1969), an apparent paradox which may be resolved by the fact that the latter's subjects were already well practiced before the experimental trials began.

Thus, even for self-paced ratings of pain, the issue of a timing artifact does not appear resolved. Examining the two functions for the 25 percent load level, the successive intervals in Trial 1 were approximately 16, 18, 24, 27, and 105 sec. while for Trial 2 the intervals between ratings were somewhat more equal, approximately 13, 19, 21, 21, and 50 sec. Since Menzer, Smith, and Caldwell (1969) note that their subjects were given practice on both self-paced and irregular rating procedures before the experimental trials were initiated, interval timing may have been learned during practice with the self-paced procedure and transferred to the irregular rating procedure during the experimental sessions. Whether this transfer phenomenon does happen or not is an empirical question which should be resolved by further research. Also, the Menzer, et al. subjects had a much lower average task duration than Lloyd, et al. (approximately 65-96 sec. vs. 120-190 sec.) for the same 25 percent load level. This suggests that, in general, the briefer the task duration, the more equal the intervals will be between successive self-paced ratings of pain. Perhaps this difference exists since it is easier to divide a short interval into equal subdivisions than a longer one. While these comments have been offered critically to evaluate the self-paced rating technique, the ability to time work duration could turn out to be an actual phenomenon representing an important parameter used to "pace" work performance.

#### Thermal Discomfort

Gagge, Stolwijk, and Saltin (1969) used rating scales to relate thermal sensation and discomfort to the level of work being performed. Four male subjects rode at work loads of 25, 50, and 75 percent max  $\dot{V}O_2$  on a bicycle ergometer under conditions of different ambient temperature (10°, 20°, and 30°C). A four-point scale of discomfort was used ranging from "comfortable" (1) to "very uncomfortable" (4). Thermal sensation was measured