Hancock, P.A., Meshkati, N., & Robertson, M.M. (1985).
Physiological reflections of mental workload. Aviation, Space, and Environmental Medicine, 56, 1110-1114.

SCIENCE NEWS NOTE

Physiological Reflections of Mental Workload

P. A. HANCOCK, Ph.D, M.Sc., N. MESHKATI, Ph.D, M.S., and M. M. ROBERTSON, M.S., B.S.

Department of Safety Science, Department of Industrial and Systems Engineering, and Human Factors Department, University of Southern California, Los Angeles, California, 90089

HANCOCK PA, MESHKATI N, ROBERTSON MM. *Physiological reflections of mental workload.* Aviat. Space Environ. Med. 1985; 56:1110-4.

Studies reporting physiological responses as reflections of mental workload are reviewed briefly. The differing measures are located in a two-dimensional space whose axes represent first, practical application and second, relevance to actual central nervous system activity as viewed from spatial and systemic congruence. Traditional methods such as those using heart rate are identified as the most practical current measures, while evoked cortical potentials emerge as superior upon the latter axis. The potential of auditory canal temperature as an optimal composite measure is explored.

ONE PARTICULARLY important question in aviation activity—as it is in the operation of many complex human-machine systems—is how hard does the pilot or operator work? If this were only physical effort then a variety of reliable and accurate techniques could be proposed (1). However, much current workload is characterized as non-physical in nature and is generally known as either mental workload or mental effort. Such a load taxes components of the active central nervous system (CNS), but as yet there are no totally reliable yardsticks against which individual response to this external stressor may be measured. The present paper reviews evidence from one tactic—variations in the physiological reaction of the performer—which has been used to investigate such load. This is by no means the only approach currently employed. Indeed, mental workload has been assessed in several differing ways, including primary task performance, secondary task performance measures, subjective estimates, and finally physiological reflections. The first three methods are not the direct concern of the present review, although informative analyses of such research are available (20,32).

Several different theoretical and experimental studies have demonstrated the relevance of physiological measures in the assessment of mental workload. It has been suggested that the interaction between the individual and the environment, particularly task-related aspects, will be reflected objectively in some physiological process. This concept has been termed *organic cost* (4). Similar reasoning has been used by Wierwille (31), who stated the underlying rationale in the following terms. As operator workload changes, involuntary variation occurs in human physiological processes; in consequence, workload may be assessed by the monitoring of the appropriate physiological system. As mental workload presumably affects the activity of the CNS, measures may variously reflect processes such as demand for increased energy, progressive degradation of the system, or homeostatic action of mechanisms designed to restore system equilibrium disturbed by such cognitive task requirements (28). However, it is difficult to distinguish between *activation* specific to the perception of the individual operator as compared with actual workload imposed. Thus, Ursin and Ursin (28) recognized that physiological methods

This manuscript was received for review in October 1984. The revised manuscript was accepted for publication in January 1985. An earlier version of this paper was presented at the International Conference on Occupational Ergonomics, Toronto, Canada, 1984.

Author Hancock is with the Departments of Safety Science and Human Factors, author Meshkati is with the Department of Industrial and Systems Engineering, and author Robertson is with the Human Factors Department at the University of Southern California, Los Angeles.

Address reprint requests to: P.A. Hancock, Ph.D., Department of Safety Science, Institute for Safety and Systems Management, University of Southern California, Los Angeles, CA 90089-0021.

do not measure the imposed load but rather they give information concerning how the individuals themselves estimate the load and in particular whether they are able to cope with it. We would suggest, however, that this is precisely the information which is required in operation of complex machine systems. Mere specification of the input load in terms of physical values of impinging stimuli would give little indication of exactly how the dynamic person-machine system would react in the applied setting.

In the present review, we are particularly concerned with the use of physiological measures in the practical work environment. Initially, and in accord with previous work (28,31), we would like to define mental workload as a reflection of purposive activity in the CNS of the sentient operator. All subsequent reference to mental workload is made with respect to this definition.

OVERVIEW

Previous reviews have adopted a largely empirical (31) or more theoretical approach (28). In this overview differing physiological methods of measuring mental workload are located in a two-dimensional space as shown in Fig. 1. The axes have been selected to represent first, on the ordinate, a practicality/impracticality scale. This is concerned solely with the question of how practical the measure is under specific working conditions including the arduous aviation environment. Considerations pertinent to placement on this axis are factors such as 1) the cost of both hardware and software to operate the system, 2) how easily is such a measure utilized by both trained and untrained individuals, and 3) how reliable is the measure particularly in the face of intervening environmental stressors which are to be expected in present day complex systems.

The second axis is constructed to represent the spatial and systemic congruence of the measure with respect to the active CNS. Spatial congruence refers to the actual spatial distance from the suggested site of mental activity. Thus measures of oculomotor

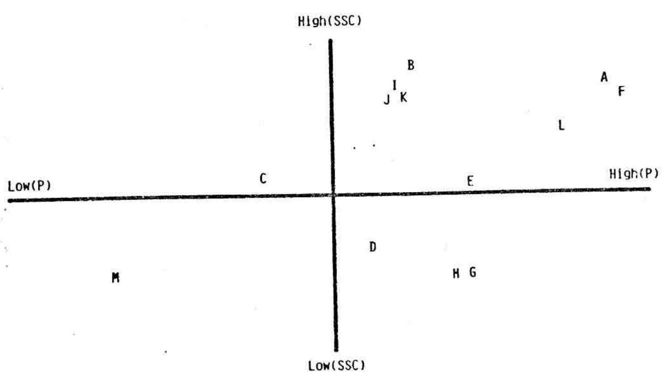

Fig. 1. Major physiological measures of mental workload located in a two-dimensional space whose axes represent 1) spatial and systemic congruence (SSC) and 2) practicality (P). Individual measures are represented by letters as follows: A = Auditory Canal Temperature (ACT); B = Event Related Potentials (ERP); C = Flicker Fusion Frequency; Critical Fusion Frequency (FFF, CFF); D = Galvanic Skin Response (GSR); E = Electrocardiogram (ECG); F = Heart Rate Variability (HRV); G = Electromyography (EMG); H = Muscle Tension; I = Electroencephalographic activity (EEG); J = Eye/Eyelid Movement; K = Pupillary Dilation; L = Respiration Analysis; M = Body Fluid Analysis.

activity score high on this particular element of the scale while alternate measures such as galvanic skin response (GSR), being somewhat spatially more remote from the CNS, rank somewhat lower. Systemic congruence refers to the nature of the interconnection of the physiological function with CNS activity. Therefore while measures such as evoked cortical potêntials score particularly highly on this component of the scale, alternate functions such as those representing cardiovascular activity score less highly. Fig. 1 represents our estimate of the location of the major current physiological measures represented within these two axes. As is immediately apparent, on the latter spatial/systemic axis the measures which emerge as the most superior are those reflecting some sort of electroencephalographic (EEG) activation. Principal among these methods are the recent developments concerning measurement of event-related brain potentials.

Event-Related Cortical Potentials

Event-related potentials (ERP's) or evoked cortical potentials (ECP's) are fluctuations in the endogenous activity of the nervous system recorded in response to environmental stimulation in association with prescribed psychological processes, or in preparation for motor activity (19). Typically, such transient voltage oscillations are recorded from the scalp using the 10/20 system of electrode placement. Repeated presentations of a physical stimulus elicit waves which are subsequently signal-averaged to reduce or eliminate random variation, to yield a stimulus locked wave or ERP. Various ERP components have been taken as reflective of information processing activity, updating system status and changes in mental workload. Workload inferences are based upon amplitude and latency elements of the elicited wave. For example, it has been noted that the amplitude of the wave observed at a latency of 200 ms following stimulus onset (P2) and the overall maximum power in the evoked response provided a metric of subjective difficulty of task performance (25).

Wisner (33) has related ERP amplitude to the attention given by the observer to a task. Other workers have indicated a correlation between amplitude and latency of visual response with fluctuation in attention during a visual vigilance task (7). One hypothesis is that the ERP reflects differences in the way the operator approaches a task and provides an indirect measure of the perception of primary task importance (2). This suggestion is similar to the notion of activation, as elaborated earlier (28), but on a more local scale. A series of studies by Wickens, Donchin, *et al.* have investigated different attributes of the ERP waveform, most notably the positive wave with a latency of approximately 300 ms (P3). In one study, an auditory task was combined with a manual tracking task. Results indicated a reduction in P3 amplitude with the addition of the auditory task compared with the no tracking control (30). Further, these investigators have shown that P3 may act as a selective index of perceptual or display processing workload (12). A second experiment studied ERP's as indicies of attention allocation. Results indicated that the P3 component of the ERP is a useful

Hancock, P.A., Meshkati, N., & Robertson, M.M. (1985).
Physiological reflections of mental workload. Aviation, Space, and Environmental Medicine, 56, 1110-1114.

metric of attention focus during discrete and continuous visual monitoring and also that the amplitude of P3 could distinguish between the two levels of workload imposed (12,30).

The ERP may be affected by certain psychological attributes. For example, it has been shown that a correctly detected stimulus at absolute threshold yields a much larger evoked potential than a more intense but irrelevant stimulus (27). Also, identical stimuli elicit different evoked responses when appearing as the target compared to appearance as some form of feedback concerning the correct identification of a target. Finally, a stimulus whose identity is known in advance produces a smaller evoked potential than its unknown and surprising counterpart. Such observations suggest that certain caution should be exercised when utilizing the ERP as a measure of workload.

The relative advantages and disadvantages of this physiological technique have been presented previously (29). Among the major advantages is the notion that ERP's represent direct reflections of the information processing activity of the operator. In addition, ERP's are multivariate measures characterized by differing latency peaks which provide a considerable amount of potential information per observation. Also, as ERP's are elicited by discrete events in the environment there is more specificity between stimulus and response compared to other more global methods. Wierwille (31) has acknowledged that the dependence of the ERP upon the operator's perceived utilities, attitude and understanding, in addition to the imposed load represents both an advantage and limitation with respect to its power as a workload assessor.

The major disadvantages of this approach include: first, the single recorded trial response contains a high noise-to-signal ratio (29). Consequently, either multiple trial recording and subsequent signal averaging, or filtering and application of analysis such as template matching is required to extract meaningful information from single observations (26). Second, the response may be contaminated by motor artifacts. Also, because of individual differences, 'calibrations' must be undertaken for each different operator. In practical terms, the technique requires considerable supporting instrumentation (e.g., computer facilities) and trained personnel for operation and interpretation. These limitations, however, represent technical barriers which are subject to constant change. Therefore, ERP's potentially represent the most promising physiological measure of mental workload for future exploitation.

Heart Rate Variability

If ERP's represent the highest scoring physiological measure on the scale of spatial and systemic congruence with respect to CNS activity, then measures pertaining to heart rate and its derivatives are currently the most practical method of assessing imposed mental workload. Among such measures, perhaps none is more thoroughly investigated than that of heart rate variability (HRV). The underlying hypothesis of the relationship between mental workload and HRV was developed primarily by Lacey (17). Essentially, the original Lacey hypothesis related the directional fractionation of cardiac activity according to the type of situation in which information occurred. This theory has been interpreted in a relatively practical manner and it has been argued that the cardiovascular system exerts some control over the bulbar inhibitory area within the brain, an area which appears to control the duration of stimulus evoked cortical activation (4). As a consequence, short term cardiac deceleration occurring both prior to and during a stimulus event could be a physiological mechanism facilitating stimulus detection. Even though Lacey's hypothesis has been disputed (3), its validity has been verified by various research groups (6,19).

Foremost among recent researchers in this area, Kalsbeek has noted a gradual suppression of heart rate irregularity due to increases in the difficulty of a task (15). In consequence, it was posited that such a measure could be used to reflect mental workload. Several empirical investigations attest to the strength of this assertion (14,16). Since such observations were first made, there have been many experimental studies in which the connection between HRV and mental workload has been observed. Detailed reviews of these efforts are available (21,31).

Measures of HRV have been assessed through the use of three major calculational approaches: 1) scoring of the heart rate data or some derivative (e.g., standard deviation of the R-R interval) 2) through the use of spectral analysis of the heart rate signal, and 3) through some combination of the first two methods. There are two advantages to such a measure—a relative and an absolute advantage. First, the absolute advantage refers to the sensitivity of the measure to change in mental workload as demonstrated in the previously cited investigations. The second advantage, inherent in the placement of this method with respect to Fig. 1, is its practical utility and relative simplicity of both administration and subsequent interpretation, when compared to alternate physiological techniques. However, it should be acknowledged that since HRV is a particularly sensitive physiological function, it is vulnerable to potential contamination from the influence of both stress and the ambient environment (13).

Tympanic Temperature

Each of the major methods as outlined above have certain strengths and weaknesses which either enhance or curtail their utility as a workload measure. These are indicated in part by their relative location within the two dimensional space of Fig. 1. However within this picture, tympanic temperature or more correctly deep auditory canal temperature (ACT) is posited as a useful composite measure which circumvents certain problems inherent in alternate approaches. This suggestion is based upon an initial empirical examination and some additional indirect observations (8,9,18,24). Following the reasoning of Ursin and Ursin (28), a load imposed upon the CNS will initiate increased activity in that structure. Accompanying such activity will be a demand for greater energy resources, directed to the appropriate areas within the CNS, in order to achieve desired performance. Data from brain scan techniques have demonstrated, as would be expected, increased metabolic activity in differing structures depending upon

Hancock, P.A., Meshkati, N., & Robertson, M.M. (1985).
Physiological reflections of mental workload. Aviation, Space, and Environmental Medicine, 56, 1110-1114.

performance required, i.e., increased activity in the motor cortex when the subject taps repetitively (18).

Utilization of energy resources may be reflected in local temperature fluctuation, although such intracranial monitoring requires a considerable investment in specialized equipment. The most suitable semi-invasive measuring site for observing such change is in the deep auditory canal and the relative merits of ACT as reflecting central temperature change in the cortex have been documented (11). It has been observed that subjects beginning work after a period of quiescence exhibit small but consistent increases in ACT (8,9). Also, subjects encountering a number of difficult calculational problems embedded in a series of simple mathematical additions show an increase in ACT (10). Finally, subjects exhibit a slight but systematic increase in ACT with activity on a mental workload, military fire command task (24).

While ACT is relatively easy to record and has in common with other physiological measures the advantage of little interference with primary task performance, it is somewhat sensitive to environmental variables, especially the fluctuation in ambient temperature. Compared with other physiological measures, the temperature signal exhibits relatively slow changes and would probably require some form of signal quickening, via the application of trend monitoring, for use in practical work conditions. Partly for this reason, the rate of change of temperature (ΔT) rather than the absolute temperature may be a superior measure and may help ameliorate the effect of individual difference inherent in all physiological parameters. It should be noted that ACT has not previously been considered as a measure of imposed load (31) and, to our knowledge, this represents the first occasion upon which auditory canal temperature (ACT) has been advocated as a physiological assessor of mental workload. Due to the potential practical benefits, ACT is the subject of continuing research with respect to both validity of measurement of central temperature change and as a reflection of human mental workload (24).

Alternate Physiological Measures

There are many alternate physiological measures of human mental workload as illustrated in Fig. 1. Since these rank progressively lower on each of the identified axes, only a brief overview of each is provided here but more detailed information on these measures is available (21,31). Among these measures there are several assessments of various forms of eye function. Flicker fusion frequency (FFF) or critical fusion frequency (CFF) is a test which determines the lowest frequency at which a flickering light is perceived as a constant stimulus by a subject. While the flickering light is of itself a stressor, its fusion frequency may be only indirectly related to mental workload, and at the same time is vulnerable to artifacts such as fatigue, learning, and metabolic rate (31).

Spatially more peripheral measures, such as Galvanic Skin Response (GSR) have been the subject of considerable interest. However, with respect to mental workload the findings have not been consistent although there is some evidence of the sensitivity of this measure to individual differences in response to a standard imposed workload (31). Electromyography (EMG) in the peripheral musculature has also been utilized as a potential measure of mental activity as was originally envisaged by Adrian and his colleagues (10). However, its sensitivity to factors other than those related to the task at hand make this measure of doubtful value.

The electroencephalogram records electrical activity of the brain through electrodes attached to the scalp. It is used extensively as a measure of mental activity and has therefore been extended into the workload realm. However, it has been pointed out that there is no simplistic relationship between the EEG record and concomitant behavioral activity and consequently, this global measure is of less value than its stimulus locked counterpart, the ERP (28). Further discussion of this measure may be found in the work of Offenloch (23). Finally, eye and eyelid movements have also been used to reflect changes in mental workload (31). However, they, as well as several alternate measures, do not provide consistent and uni-directional response to workload change and are in consequence of dubious utility. Gaume and White (5) investigated the relationship between mental workload and respiration. They found that integrated respiration volume (IRV) was a relatively sensitive indicator of mental workload level. In addition to the methods discussed in the present section, there have been attempts to use various alternate physiological functions and indeed it may be suggested that few physiological functions could not be used as workload measures. The utility of each of the above and alternate methods as shown in Fig. 1 is discussed in the work of Meshkati (21) and others (20,22).

CONCLUSION

In this brief review, various physiological responses which provide reflections of human mental workload have been examined. These measures have been located along two axes pertaining first, to the current practicality of the measure and second, to the spatial and systemic congruence of the measure with respect to the active central nervous system. While the suggested superior measures on each axis, i.e., heart rate variability and event related potentials, have been reviewed in some detail, alternate physiological measures have received somewhat less attention. It has been suggested that temperature change, as monitored in the deep auditory meatus, may provide a valid composite which in part palliates those problems experienced in alternate methods. While the use of individual physiological measures is possible, it may be more appropriate to combine a number of such measures into a composite physiological test battery (22).

While physiological measures represent only one strategy in the approach to the problems on mental workload, they do contain the advantage of relative non-interference with the primary task at hand. Therefore, they hold a particular ascendency over other invasive measures, especially in circumstances where the operator cannot tolerate forms of external distraction (e.g., low-altitude night aerial navigation). Some investigators have advocated a multiple battery approach, using representative elements from each of the four major workload assessment methods as noted earlier. In such an approach, however, the advantage of non-interference is sacrificed for the greater specificity of workload information obtained. In the operation of dynamic person-machine systems, the rate limiting factor is frequently operator capability. Consequently, the limit of such abilities and the operator's approaching of his own individual load tolerance are of particular importance with respect to safety and efficiency of overall action. Physiological reflections of mental workload hold the potential to be among the most efficacious assessors of such an important performance parameter.

ACKNOWLEDGMENTS
The authors would like to thank the editor and the unknown reviewers for their helpful comments on this paper.

REFERENCES

1. Astrand P, Rohdahl K. A textbook of work physiology. New York: McGraw Hill, 1977.
2. Defayolle M, Dinand JP, Gentil MT. Averaged evoked potentials in relation to attitude, mental load and intelligence. In: Singleton WT, Fox JG, Whitfield D, eds. Measurement of man at work. London: Taylor and Francis, 1973.
3. Elliott R. The significance of heart rate for behavior: a critique of Lacey's hypothesis. *J. Pers. Soc. Psychol.* 1972; 22:398–409.
4. Firth PA. Psychological factors influencing the relationship between cardiac arrhythmia and mental load. *Ergonomics.* 1973; 16:5–16.
5. Gaume JG, White RT. Mental workload assessment, III. Laboratory evaluation of one subjective and two physiological measures of mental workload. Long Beach, CA: McDonnell Douglas Corp. Report MDC-J7024/01, 1975.
6. Graham FK, Clifton RK. Heart-rate change as a component of orienting response. *Psychol. Bull.* 1966; 65:305–20
7. Groll-Knapp E. Evoked potentials and behavior. In: Singleton WT, Fox JG, Whitfield D, eds. Measurement of man at work. London: Taylor and Francis, 1973.
8. Hancock PA. The effect of an induced selective increase in head temperature upon performance of a simple mental task. *Hum. Factors.* 1983; 25:441–8.
9. Hancock PA. An endogenous metric for the control of perception of brief temporal intervals. *Ann. N.Y. Acad. Sci.* 1984; 423:594–96
10. Hancock PA, Brainard DM. Tympanic temperature: a non-invasive physiological measure of workload. Massachusetts: Endeco Inc: Technical Report, 1981.
11. Hancock PA, Dirkin GR. Central and peripheral visual choice reaction time under conditions of induced cortical hyperthermia. *Percept. Mot. Skill* 1982; 54:395–402.
12. Israel J, Wickens CD, Donchin E. The event-related brain potential as a selective index of display load. *Proc. Hum. Factors Soc.* 1979; 23:558–62.
13. Kalsbeek JWH. Standards of acceptable load in ATC tasks. *Ergonomics.* 1971; 14:641–50.
14. Kalsbeek JWH. Do you believe in sinus arrhythmia? *Ergonomics.* 1973; 16:99–104.
15. Kalsbeek JWH, Ettema J. Scored regularity of the heart rate pattern and the measurement of perceptual or mental load. *Ergonomics.* 1963; 6:306.
16. Kalsbeek JWH, Sykes RN. Objective measurement of mental load. *Acta Psychol.* 1967; 27:253–61.
17. Lacey JI. Somatic response patterning and stress: some revisions of activation theory. In: Appley MH, Trumball R, eds. Psychological stress: issues in research. Appleton-Century-Crofts: New York, 1967.
18. Lassen NA, Ingvar DH, Skinhoj E. Brain function and blood flow. *Sci. Am.* 1976; 239:62–71.
19. Martin I, Venables PH. Techniques in psychophysiology. New York: Wiley, 1980.
20. Moray N, ed. Mental workload, its theory and measurement. New York: Plenum Press, 1979.
21. Meshkati N. A conceptual model for the assessment of mental workload based upon individual decision styles. [Dissertation]. University of Southern California, 1983.
22. O'Donnell RD. Development of a neurophysological test battery for workload assessment in the U.S. Air Force. *Proc. Int. Conf. Cyber. Soc.* IEEE, 1981.
23. Offenloch K. Neurophysiological assessment of functional states of the brain. *Proc. AGARD Conf. Workload.* AGARD-CP-216, 1977.
24. Robertson MM, Hendrick H, Hancock PA. Individual response to a computer generated mental workload task as a function of cognitive complexity. Paper presented at the International Conference on Occupational Ergonomics, Toronto, Canada, 1984.
25. Spyker DA, Stackhouse SP, Khalafalla AS, McLace RC. Development techniques for measuring pilot workload. Tech Rep. NASA CR-1888, 1971.
26. Squires KC, Donchin E. Beyond averaging: the use of discriminant functions to recognize event related potentials elicited by single auditory stimuli. *EEG Clin Neurophysiol.* 1976; 41:449–59.
27. Sutton S, Tueting P. The sensitivity of the evoked potential to psychological variables. In: Venables PH, Christie MJ, eds. Research in psychophysiology. New York: Wiley, 1975.
28. Ursin H, Ursin R. Physiological indicators of mental workload. In: Moray N, ed. Mental workload, its theory and measurement. New York: Plenum Press, 1979.
29. Wickens CD. Measures of workload, stress and secondary tasks. In: Moray N, ed. Mental workload, its theory and measurement. New York: Plenum Press, 1979.
30. Wickens CD, Israel JB, Donchin E. The event-related cortical potential as an index of task workload. *Proc. Hum. Factors Soc.* 1977; 21:282–6.
31. Wierwille WW. Physiological measures of air crew mental workload. *Hum. Factors.* 1979; 21:575–93.
32. Williges RC, Wierwille WW. Behavioral measures of air crew mental work load. *Hum. Factors.* 1979; 21:549–74.
33. Wisner A. Electrophysiological measures for tasks of low energy expenditure. In: Singleton WT, Fox JG, Whitfield D, eds. Measurement of man at work. London: Taylor and Francis, 1973.