N88-23371

140789
18P.

AN OVERVIEW OF CURRENT APPROACHES AND FUTURE CHALLENGES IN PHYSIOLOGICAL MONITORING

Richard L. Horst
ARD Corporation
Columbia, MD

ABSTRACT

Sufficient evidence exists from laboratory studies to suggest that physiological measures can be useful as an adjunct to behavioral and subjective measures of human performance and capabilities. Thus it is reasonable to address the conceptual and engineering challenges that arise in applying this technology in operational settings. The present paper will attempt to identify such application-oriented issues and to provide an overview of the state-of-this-art. Issues to be reviewed will include the advantages and disadvantages of constructs such as mental states, the need for physiological measures of performance, areas of application for physiological measures in operational settings, which measures appear to be most useful, problem areas that arise in the use of these measures in operational settings, and directions for future development.

INTRODUCTION

Prospects for the routine use of physiological monitoring in operational settings are becoming more favorable. This situation is due in part to advances in recording technology, in part to research results that suggest the usefulness of physiological data, and in part to an increasingly critical perceived need for information about the status of the human operator in complex man-machine systems.

One can sometimes gain an impression of the state of one's art by the criticism it receives during informal exchanges. Not many years ago, those of us involved in psychophysiological research, and in particular scalp-recorded brain-wave measurement, were frequently asked to endure two comments:

> "Surface recordings provide only a gross indication of brain function. It's like putting an electrode on the outside of a computer and trying to infer the processes going on inside."

and

> "How can you interpret these field potential phenomena without understanding the underlying mechanisms, if not the underlying physiology?"

Perhaps it is the company one keeps, but lately other comments have been heard more frequently:

> "You can't have electrode wires dangling from a pilot in the cockpit."

> "Operators will never accept having their physiology monitored. It takes too long to hook them up. It's too messy. Besides, pilots will be afraid that you'll turn up some arrhythmia that could ground them."

> "What do you do with all the electrical artifacts that are likely to show up in operational settings? In the laboratory you can reject contaminated data and keep collecting until you get enough clean data. In the field you will not have that luxury."

> "There is no one-to-one relationship between (fill in your favorite physiological sign) and performance. You would have to know a lot about overt behavior in order to interpret concurrently recorded physiological measures. And if you have the behavioral measures, why do you need the physiological?"

Thus, the issues of concern seem to be changing, from questioning the basic value of the measures to questioning how one implements them in applied settings. There is no question that much basic research and theorizing remain to be done in this field. We don't yet have a good understanding of the functional significance of many psychophysiological phenomena. But, as funding permits, progress is being made and physiological measures are proving to be valuable adjuncts to behavioral and subjective measures in the assessment of human performance (see Ref. 1 for a recent broad survey of this field). For this purpose, derived measures of physiological signals can be useful as dependent measures, regardless of how poorly we understand the underlying physiology. A thorough understanding of source generator loci and cellular mechanisms would, no doubt, enhance the interpretive power of these measures; but as long as they vary systematically with experimental manipulations, these indices can be used, as are behavioral and subjective measures, in the monitoring, prediction, and diagnosis of performance.

Corresponding to this shift in the concerns of critics, one notices an attitudinal change among practitioners. For years, basic researchers took a rather cavalier approach -- that their role was to demonstrate the value of psychophysiological measures of performance and to uncover the relationships between these measures and conceptual information-processing constructs. Problems related to the transition of this technology to applied task environments and the implementation of these measures in the field could be left to "the engineers." Now, one finds considerable interest, among both researchers and funding agencies (one can speculate about the causal relationships here), in beginning to address these deferred "engineering" problems. Impetus has been provided by advances in a number of enabling technologies -- micro-electronics, signal processing, wireless communications, display technology, and artificial intelligence (AI). Consequently, laboratory work is being conducted with an eye towards task scenarios and measurement protocols that could, with modification, be used in the field. More research, both basic and applied, is being conducted in simulators.

All of this represents progress and suggests the need to look closely at the realistic prospects for applying physiological measures in operational settings. The remainder of this paper will provide a necessarily brief overview of some of these prospects, the approaches that are currently being pursued, the state-of-the-art, and recommendations for future directions in research and development. One theme, which corresponds to the topic of this workshop, will be the prospects for quantifying operator mental states.

MENTAL STATE ESTIMATION

It is interesting that in the conceptual plans for such next-generation systems as those involving Super-Cockpits, one sees a recognition of the fact that operator mental status is something the system should measure and to which it must adapt. No doubt, this design goal follows from the recognition that, under some operational scenarios, the human operator could be the limiting factor for successful mission completion. These systems will be capable of presenting more information than even a fully functional human can process, and some of the threats faced in the operational environment, e.g., high G load or chemical/biological/radialogical (CBR) agents, could disable the operator without fatally impairing system hardware and software. Moreover, these systems are expected to have sufficient automated subsystems and artificial intelligence that the system could aid an overburdened operator or, to some extent, take over for an impaired operator.

Certainly, therefore, the ability to assess the functional mental status of the human operator is of critical importance in these systems, and would be useful to the designers of many less exotic systems. But how far can we take this concept? Can one conceptualize functional mental status in terms of a finite number of discrete mental states? Is there some value to being able to classify the human operator from moment to moment as being in a state of high or low workload, fatigue, boredom, confusion, stress, or any of the numerous other explanatory constructs that we invoke, even informally, in interpreting our data or in designing our man-machine interfaces?

Typically, these constructs are operationally defined in terms of experimental variables. Beyond that, it is not yet clear whether such discrete states exist, or with what taxonomy they should be classified. Operator effectiveness is ultimately defined in terms of behavioral output. However, there seems to be both diagnostic and prescriptive value in attempting to develop such a taxonomy of mental constructs, rather than focusing just on observable task performance. For example, task performance may deteriorate for a wide variety of reasons. An operator may miss an alarm signal either because he was cognitively overloaded or because he was bored and not sufficiently vigilant. A system designer, or co-pilot, would take different remedial actions, depending on which of these "states" led to the degradation in performance. Furthermore, many task environments allow the human operator to function with some spare capacity such that, to some extent, increased task demands can be met with increased effort in order to maintain behavioral output at a relatively constant level. In such situations, mental state indices may predict susceptibility to an impending deterioration in performance, should task demands increase still further. Finally, when task demands are low, there may be little behavioral output from which one can gauge the status of the operator. A sense of the operator's mental state in such situations could be used to infer whether or not such lack of responding was appropriate and the extent to which the operator is prepared to respond appropriately should conditions change. Therefore, the diagnostic and, hopefully, prescriptive value of mental state constructs are somewhat akin to that of clinical syndromes. Analogous to the different treatments which may be prescribed depending on a clinical diagnosis, inferences about the mental states which underlie an observed performance deficit may suggest alternative design or operational "treatments."

The danger in using mental state conceptualizations to explain data, of course, lies in our tendency to think that if we can label something, we have understood it. Terms like "boredom" may not imply the same "syndrome" to everyone. Therefore, until we have sufficient data to define what are the distinguishing features and performance-related consequences of "boredom," it is imperative that we continue to operationally define our use of such terms.

THE VALUE OF PHYSIOLOGICAL MEASURES

Regardless of the stock one puts in the explanatory power of mental states, it follows from the above discussion that it would be unwise to evaluate and predict an operator's ability to perform solely from observing behavior on a primary task. Performance on secondary tasks can be instructive for measuring the processing capacity entailed by a primary task. However, with this approach it is difficult to ensure that the operator always gives mental priority to the primary task, the results may be of questionable validity if used to generalize to situations in which the primary task is performed alone, and incompatibilities between the behavioral responses required by the two tasks may make it difficult to draw inferences about the demands placed on perceptual or decision-making processes. Moreover, the sort of contrived secondary tasks that have often been used in laboratory studies are clearly not acceptable in operational settings, so secondary task measures must be found among the activities that the operator is doing in the course of normal operations.

Simply asking the operator for subjective ratings of his perceived state is often useful, but is also fraught with difficulties. The operator may not realize that his environmentally-defined workload is high when, in fact, it is. Furthermore, such subjective ratings tend to be unreliable when administered in operational settings while the operator is simultaneously trying to maintain task performance, and the mere act of completing the rating itself, of course, constitutes an additional task burden on the operator.

For these reasons, there is considerable appeal to the prospects of gaining additional information about the functional status of operators from their physiological signs. As discussed later in this paper, much evidence now suggests that, if interpreted in conjunction with behavioral and subjective measures, physiological indices offer the possibility of objectively inferring, not only the general physical fitness to perform, but also the cognitive status of an operator. Physiological measures can often be used to confirm the conclusions derived from behavioral or subjective measures. There are also instances in the literature of physiological measures providing complementary information regarding cognitive activity to that which is available from behavioral measures.

While there is a certain intuitive appeal to the objectivity and non-intrusiveness afforded by physiological measures of mental processes, the possible limitations of this technology have been pointed out by a number of critics. Johnson (Ref. 2) has listed several typical concerns:

- Most research studies have used performance changes to interpret physiological changes; it is the inverse problem, using physiological indices to predict performance, that is of interest in operational settings, and most attempts to take this approach have been disappointing.

- There are not specific physiological response patterns associated with specific behaviors or specific states; task difficulty plays an important mediating role.

- There are large individual differences in physiological responses; response differences due to individual response stereotypy tend to be larger than differences due to situational response stereotypy.

Zacharias (Ref. 3) has likewise faulted most physiological work for failing to take account of the effects of task difficulty on the measures of interest. He also points out that while attempts to more fully characterize physiological status by creating a vector of physiological indices may provide increased correlations with, for example, measures of workload, there can actually be a reduction in the statistical significance of such correlations, as "an increasing number of noisy physiologic indicators are included in the actuation vector."

While these criticisms are well-taken and must be addressed by those wishing to use physiological measures of performance, they pose no insurmountable problems for the knowledgeable application of physiological monitoring technology. It is possible to deal with, and in fact take advantage of, the manner in which physiological indices reflect task difficulty (see, for example, Samaras'[1] paper in the present Proceedings). The irrefutable fact that individual differences exist, may likewise be turned to our advantage. In most operational settings we are dealing with highly trained operators, and it is technologically possible to customize the parameters of a monitoring system for the individual operator. Finally, the question of whether or not unique configurations of physiological patterns can be associated with particular mental states may be moot, if one assumes that interpretations can be based on changes in physiological indices viewed in conjunction with changes in operator behavior or system performance. In other words, one rarely would be faced with the need to classify operator state in an absolute sense. The more frequent, and more manageable, challenge would be to classify changes in state or functional status, in relative terms, with reference to task performance and other behavioral data.

AREAS OF APPLICATION

Physiological measures can be useful in operational settings for a variety of purposes. Other papers in this session have presented some specific operational settings of interest. Most uses can be seen to fall into one of the following categories:

System Design. Reducing operator workload and drawing an operator's attention to certain task-related stimuli are often design goals. To the extent that physiological measures are reliable indices of these mental constructs, they can be used to make design decisions. For this group of applications, recording in facilities that simulate the operational environment is useful, data analysis can be done off-line, and, consequently, we have the luxury of dealing with measures based on derived indices such as average waveforms. Applications of this sort would include:

- Choosing among alternative hardware or software.

- Choosing among alternative procedures.

---

[1] Samaras, George M: Towards a Mathematical Formalism of Performance, Task Difficulty, and Activation. NASA CP 2504, 1988, pp. 43-55

- Assessing the fidelity of simulation.

- Use as a debriefing tool, to probe operators with additional questions, after-the-fact, about the times during a recorded scenario when the physiological signs suggested, for example, that the operator was stressed or distracted.

On-line, Real-time Applications. To the extent that physiological indices of performance can be extracted on one or a few trials (i.e., from single epoch recordings), and it is feasible to derive these indices in real-time in the operational setting, they would be useful in closed-loop man-machine systems. In general, this group of applications would involve the feedback of physiological information from the operator to the machine with which he is interacting, so that decision-making algorithms that reside there can modify the operator's task or displays accordingly. This group of applications is perhaps most demanding, because of the need for real-time turnaround of the measures of interest. Applications of this sort would include:

- Assessing the general state of the operator, to determine whether he is fit to be "in the loop" at all.

- Dynamically allocating tasks between the human operator and onboard AI, depending on workload.

- Checking whether the operator attended to events that the onboard AI flagged as significant, as well as detecting instances in which the operator realizes he made an error, so that he has an opportunity to correct himself.

Personnel Selection and Training. To the extent that physiological measures reflect cognitive processes for which there are significant individual differences, these measures may prove useful for selecting personnel and monitoring the progress of an individual's training. The challenge here is to define measures that are predictive of future performance. As with system design applications, we would frequently be able to process the recorded data off-line and deal with derived measures, without the constraints of real-time turnaround. Some applications of this type include:

- Staffing high workload tasks or environments with individuals who are well-suited to handle them.

- Channeling personnel into jobs that take advantage of their cognitive styles.

- Determining skills in an individual's training program that remain to be mastered by identifying the aspects of a task that cause high workload.

THE MOST PROMISING PHYSIOLOGICAL MEASURES

The research literature provides considerable evidence to suggest that a number of physiological measures will be useful for the applications mentioned above. It is beyond the scope of this paper to attempt a comprehensive review of this literature. However, in the present section a cursory overview is offered, to provide some indication of which indices of central and peripheral sweats, and moves about in performing his duties. In addition, physiologically generated artifacts, most often from eye movements and blinks, or skeletal muscle activity, can likewise contaminate the recordings. While useful recordings of EEG have been reported in-flight (e.g., Refs. 10, 11), significant engineering advances are required in electrode application, signal processing, and artifact rejection before such recordings could be used routinely.

Event-related Potentials (ERPs). ERPs are also voltage fluctuations recorded from the scalp, but those which are time-locked to events, usually external stimuli. Transient ERPs are characterized by the amplitude, latency from stimulus onset, and scalp distribution of the various component peaks in the waveform. The stimulus-locked brain activity is typically examined after signal averaging over numerous presentations of the same event, although single trial analysis techniques are an active area of investigation. ERP recordings in operational settings are subject to the same technical constraints as those of ongoing EEG.

Various ERP components have been shown to vary reliably with cognitive processes (see review in Ref. 12), including selective attention (e.g., Ref. 13), expectancy (e.g., Ref. 14), discrimination processes (e.g., Ref. 15) and response preparation (e.g., Ref. 16). In contrast to the findings regarding ongoing EEG, there is a body of research that has shown very encouraging relationships between ERP indices and workload. This work, by Donchin, Wickens, and colleagues, is reviewed in the Munson,[2] et al. paper in the present Proceedings. There is evidence that ERPs may be used to reveal systematic cognitive effects in addition to those which are apparent from behavioral measures alone. For example, P300 latency has been shown to vary with only a subset of the manipulations that affect overt reaction time, suggesting that the timing of P300 indexes the completion of stimulus evaluation processes, independent of response selection processes (e.g., Ref. 17). In certain situations, P300 amplitude appears to be a reflection of subjective probability, whereas overt choice behavior may be influenced by additional variables, for example those which affect the willingness to take risks (e.g., Ref. 18).

Steady-state ERPs are recorded in response to a rapidly oscillating stimulus, usually a light or sound. They are usually quantified in terms of amplitude and phase delay at the frequency of stimulation, and can be calculated after only several seconds of stimulation. Steady-state ERPs elicited by rapid, periodic stimulation by a checkerboard have also been reported to reflect workload when the checkerboard was presented concurrently with task performance (e.g., Ref. 19). This result is surprising, given that steady-state responses had been previously thought to reflect strictly sensory processes. The effect needs to be further examined to rule out the possibility that peripheral changes in the visual system, such as accomodation, could be varying with task difficulty and thus mediating the changes in the steady-state response.

Electrooculography (EOG). EOG recordings are derived from electrodes on the face near the eyes and can be used to monitor eye movements, eye blinks, and,

---

[2] Munson, Robert C.; Horst, Richard L.; and Mahaffey, David L.: Primary TASK ERPs Related to Different Aspects of Information Processing. NASA CP 2504, pp. 163-178.

nervous system activity, of the many that can be recorded non-invasively from behaving humans, appear most promising for near-term application.

Although there is considerable overlap in the measures that appear useful for different kinds of applications, a distinction should be made between the use of physiological measures to indicate the basic fitness of the operator to perform his tasks and the use of physiological measures to infer cognitive status. The former applications entail the monitoring of vital signs to indicate relatively gross impairments in physical well-being — e.g., G-induced loss of consciousness or gray-out, exposure to CBR agents, motion sickness, heat stress, traumatic injury, heart attack. The latter applications entail the analysis of more subtle physiological changes related to task performance, so as to infer mental states such as high workload, fatigue, or inattention.

The following overview focuses on those measures with a demonstrated relationship to operationally defined manipulations of workload, stress, fatigue or boredom. While most of these relationships have been demonstrated in laboratory settings with non-real-time processing of the data, some have been recorded successfully in operational settings and all hold at least the promise of being feasible to derive in real-time. Typical quantitative measures that are derived from each physiological sign are presented, technical problems in recording these measures in operational settings are discussed, and examples of the evidence relating these measures to the psychological constructs of interest are mentioned. More extensive discussions of the prospects for using physiological measures in operational settings may be found in O'Donnell (Ref. 4) and Gomer (Ref. 5).

Electroencephalography (EEG). The EEG consists of voltage fluctuations recorded from two or more sites on the scalp, Ongoing EEG is usually quantified in terms of its frequency composition and amplitude asymmetries. Other measures, such as the coherence between the activity recorded at various pairs of scalp sites, also appear to be useful (e.g., Ref. 6).

Changes in the predominant frequencies in the EEG with levels of arousal and activation have been known for some time (e.g., Refs 7, 8). An alert person performing an engaging task shows predominantly low amplitude, fast frequency (beta) activity. An awake, but less alert, person shows an increased incidence of high amplitude, alpha (8-12 Hz) activity. With the onset of drowsiness, slower frequency theta (4-7 Hz) activity enters the spectrum and in the early stages of sleep, very high amplitude, slow (1-3 Hz), delta waves predominate. It is unlikely in operational settings that operators would lapse into deeper, so-called "paradoxical," stages of sleep. The generalized effect of stress, activation or arousal is, therefore, a shift towards the faster frequencies, often with an abrupt blocking of the alpha rhythm (e.g., Refs. 8, 9). Fatigue and boredom generally shift the spectrum in the other direction, towards the lower frequencies. Derived measures of ongoing EEG have not yet proven to be reliable indicators of workload.

Aside from the general problems of isolating the physiological recordings from environmental sources of electrical noise and deriving the measures of interest in near real-time, there are several technical problems in recording EEG and related measures in operational settings. Movement of the electrodes relative to the scalp causes severe electrical artifacts, and it is difficult to ensure firm contact in environments where the operator wears a helmet, to a limited extent, direction of gaze and eye closure. The EOG reflects changes in the electric dipole formed between the cornea and the retina. While these potentials interfere with scalp-recorded electrophysiological measures such as EEG and ERPs, measures derived from the EOG itself have been shown to reflect operators' cognitive state.

Blink rate increases reflect the deterioration in attention and performance which occur over a prolonged task (e.g., Refs. 20, 21). Additionally, blink durations have been shown to increase with time on task (Ref. 22). Thus, increases in both blink rate and duration may indicate fatigue or lack of vigilance. As workload increases, blink rates decrease and the latency of the blink, after presentation of the stimulus of interest, increases (Ref. 23). Moreover, blinks during visual tasks were found to be of shorter duration than those in auditory tasks (Ref. 22). The pattern of these results are consistent with the notion that as visual information processing demands increase, eye blinks reflect the brain's attempt to take in more visual input.

Blinks are robust and easy to record, because they are of relatively high amplitude and predictable waveshape. Measures of blink frequency and latency should, therefore, be feasible even in somewhat noisy environments. Measures of blink duration will, of course, require relatively noise-free signals.

Eye Position and Pupil Dilation. Eye movements and fixations, and pupil dilation, are usually detected by photo-optical techniques and, therefore, are measures that can be gathered without sensors that touch the subject. Eye position is inferred from corneal reflectance and is usually quantified in terms of direction of gaze and dwell times as the eye scans the environment. Pupil size is measured in millimeters.

Dwell time on various displays on an instrument panel has been shown to vary systematically with workload (Ref. 24). Both tonic levels of pupil size over long durations of task performance and phasic responses elicited by task-relevant stimuli have been shown to be sensitive to cognitive variables. Tonic dilations seem to be a reliable index of activation and arousal (e.g., Ref. 9). In addition, consistent phasic increases in pupil dilation have been associated with increases in task difficulty and workload (e.g., Refs. 25, 26).

Because both these indices are dependent on maintaining a beam of light on the cornea, they are limited to environments, such as fixed-base simulators, in which there is minimal head movement by the operator. Eye trackers are becoming more sophisticated, but head movements beyond about one cubic foot take the eye out of range of the presently available photo-sensors. It is likewise difficult to maintain a fix on the eye in a high-vibration environment. Further confounds can be introduced by the fact that pupil size is responsive to non-specific factors such as ambient illumination, color, and depth of the visual field, which are difficult to control in operational settings.

Electrocardiography (ECG). ECGs are a widely used, easily recorded index of cardiovascular activity that is obtained from a two- or three-electrode array on the body. The ECG signal may be analyzed in terms of its basic timing (heart rate or period) or its morphology (e.g., amplitude of the T-wave). Derived measures from the ECG, given the detection of the R-wave as the basic datum, include first-order measures such as rate per unit time and change in heart period across beats. Second-order analysis may include rate-of-change measures, maximum and minimum beat-to-beat periods within an epoch, and methods based on time-series analysis of the beat-to-beat intervals.

While heart rate has been shown to generally increase with stress (e.g., Ref. 27) and activation (see review in Ref. 9), the heart rate response to stimuli in a task environment is more often characterized by a complex pattern of deceleration and acceleration. The results of numerous (but not all) relevant studies are consistent with a hypothesis put forth by Lacey (Ref. 28), that heart rate deceleration reflects a receptivity to external stimulation whereas accelerations occur if the situation is found, after initial attention, to warrant an increase in energy release. Heart rate increases during periods of increased workload, for example during take-offs and landings, have been reported (e.g., Refs. 29, 30) but others have not found heart rate to be sensitive to the cognitive workload of simulated flight (Ref. 26).

More consistent relationships with workload have been reported for heart-rate variability. The general finding has been that, with increased attention and workload, heart-rate variability decreases (e.g., Refs. 31, 32). The most frequently used technique to reveal this workload effect has been a spectral analysis of the beat-to-beat time interval data with a focus on the power in the 0.1 Hz band (e.g., Ref. 33). Of particular interest has been the component of heart-rate variability related to respiratory sinus arrhythmia. because of the many influences on the beat-to-beat regularity of the heart, this one reflects mediation by the central nervous system. An approach to quantifying sinus arrhythmia, which makes fewer assumptions about the statistical properties (i.e., stationarity) of the data than those based on spectral analysis, is that of vagal tone. Porges[3] (see Ref. 34 and paper in this Proceedings) has developed a moving polynomial filter technique that removes the slowly shifting baseline from the inter-beat interval data over time in order to reveal the faster oscillations due to respiratory sinus arrhythmia. In the few instances in which this "vagal tone" measure has been compared to the measure based on power in the 0.1 Hz band, vagal tone has proven to be the more sensitive indicator of the experimental manipulations (Ref. 35).

Heart rate measures have been successfully recorded under extremely demanding conditions (e.g., Refs. 36, 37, 38).

Respiration. A number of techniques have been proposed for measurement of the basic respiratory signal. As a class, girth measurements of the thorax and/or the abdomen using mercury-in-silastic tubing strain gauges are simple, non-invasive, and reliable. If possible, both thoracic and abdominal components of the respiratory motion should be monitored, since it is possible to derive an adequate measure of respiratory volume from the combined signals. The principal measures are respiratory rate, average volume (if composite), and parameters related to the timing of inspiration, inspiratory pause, expiration, and expiratory pause. Tidal volume, the volume of air expired, can be sensed by thermistors mounted unobtrusively in an oxygen mask. Minute volume may vary independently of tidal volume and can be measured in the same way.

---

[3] Porges, Stephen W.: Vagal Tone as an Index of Mental State. NASA CP 2504, pp. 57-64

Respiration measures deserve more attention than they have received (e.g. Ref. 39) for detecting operator incapacity. There is also some indication that respiration becomes more shallow, regular and rapid with increased workload (Ref. 40).

Electromyography (EMG). EMG recordings from surface electrodes can be used to detect muscle tone or movement mediated by selected muscle groups, if they can be recorded without contamination by task-related movements. Several sites have been suggested as indicating overall tension levels, particularly forehead or masseter muscle placements. Since the signal is a complex, irregular one, the preferred strategy for determining general tension levels is to integrate the primary signal over a relatively short time constant, typically between 0.1 and 0.5 seconds, and to subsequently analyze only this average measure. The measures typically derived from the average muscle tension level are mean level, variance of the level, and minimum and maximum level for each epoch. If appropriate, further measures such as the number of increases above a criterion level can be obtained.

Muscle tension increases with arousal, stress and activation (e.g. Refs. 9, 41) and increased EMG activity is associated with the onset of fatigue. Several studies have reported relationships between increased EMG activity and increased workload or task difficulty (e.g., Refs. 42, 43), but it is as yet unclear as to how sensitive EMG is as an index of small changes in workload.

Other Measures of Interest. A number of other physiological measures deserve "honorable mention," either because they appear to be worthwhile indicants of cognitive status, but without the near-term prospects for application in the field, or because they appear to be related to cognition in only a general sense:

o Ongoing and stimulus-locked measures based on magnetoencephalography recordings are particularly promising because the sensor does not touch the subject's body and because inferences can often be made about the depth from which activity arises. Evoked magnetic fields have been correlated with attention and subjective probability in a paradigm similar to that used for ERP studies of P300 (Ref. 44). However, the sensors now in use must be supercooled with a large container of liquid helium and the subject must maintain a posture which keeps his orientation and distance from the sensor constant.

o Blood pressure and blood flow can provide useful information about cardiovascular status which, to some extent, complements that available from heart rate and heart rate variability. However, methods for recording these indices non-invasively have not yet reached the point that they would be useful in an electrically noisy, high vibration environment, or one in which the operator had to be free to move significantly.

o Advances are being made in the sensor technology for monitoring body temperature, with the development of miniaturized telemetry systems that can be swallowed as a "pill" and used to monitor core temperature as it passes through the gut, and with the development of improved skin temperature sensors. This technology promises to be of use in environments where heat stress is a threat, and phasic temperature changes have been related to mental workload (e.g., Ref. 45) as well as physical workload.

- Measures of skin resistance and skin conductance are relatively easy to record, and have some value for indicating phasic changes in arousal and stress, but they have yet to prove themselves as specific enough to be of utility for inferring cognitive states.

PROBLEM AREAS

There is no question that significant technical problems remain to be solved before physiological monitoring technology will come into widespread use in operational settings. But it is also apparent that recent technological developments offer new possibilities for solving many of these problems and that researchers, and funding agencies, are only now turning their attention towards these prospects. Some areas of concern are the following:

Instrumentation and Operator Acceptance. The operator's reluctance to be instrumented is an often-mentioned impediment to implementation of physiological monitoring in operational settings. Operators find conventional recording paraphernalia cumbersome and obtrusive. It is time-consuming to have electrodes pasted on and removed. They are also threatened by the possibility that in submitting to recordings, an unanticipated medical problem may be detected that could call into question their eligibility. When faced with the prospects of closed-loop decision-making, operators are reluctant to relinguish their control of a system to automated subsystems.

As recording instrumentation becomes more miniaturized, some of these objections will disappear. There are now several "pocket-size" amplifier/ recording systems available for ambulatory monitoring (e.g. the SSPIDR, see Banta's paper in this Proceedings). On-board storage of physiological data is now achieved with either cassette tape or solid-state memories. Optical disk media may soon provide still further storage capacity. Telemetry systems are likewise becoming smaller and more sophisticated. "Paste-less" electrodes have been a possibility for some time, but require further refinement. Integrating electrodes and amplifiers into helmets and uniforms remains a challenge, but is being addressed by several groups. The palatability of using physiological measures in closed-loop control systems will be increased by giving the operator the ability to override the decisions reached by the on-board decision-making algorithms, and by introducing this technology as an open-loop "aid" to the operator until the decision rules mature to the point that they warrant the operator's confidence. As for the objections which can't be addressed with instrumentation, one suspects that as the value of physiological measures becomes more apparent and the safety implications of not having them is more widely recognized, these problems will largely take care of themselves.

Safety issues. Any tethering of the pilot to recording equipment must be done in a way that does not distract or impede him from performing his duties. In some environments, such as fighter aircraft where the aircrew must be able to eject if necessary, this requirement dictates a telemetry system for transmitting the amplified physiological signals to on-board or remote processing equipment or an entirely portable physiological recording system that can be carried on the operator's person (e.g., Ref. 46). Furthermore, the recording equipment must be electrically integrated with the other equipment with which the operator interacts, so that there is no shock hazard when he touches the control stick or instrument panel.

Here again, advances in micro-electronics are allowing increased miniaturization, and thus portability, of amplifiers, storage media and telemetry systems. Amplifiers can be designed with fail-safe features to protect the subject against internal shorts in the circuitry, and the possibility of such failures can be minimized by "hardening" physiological recording equipment, using the same methods that are used for other on-board electronic instrumentation, for use in even high-vibration environments. If telemetry is used, it must be accomplished with a technique or in a frequency range that does not interfere with other on-board equipment. Ensuring against shock hazard involves issues of electrical grounding that can usually be readily solved with cooperation from system engineers.

Artifact Rejection and Compensation. There are two sets of issues regarding contamination of recordings by artifact — one involving electrical artifacts from the environment and the other involving physiological artifacts from the subject himself. Most operational settings are electrically noisy environments, so aside from the above safety issues, appropriate shielding and grounding must be implemented in order to get clean physiological recordings. Miniaturization of amplifier electronics and efforts to integrate this circuitry into helmets and suits, offers the prospects of placing the amplifier circuitry on or in close proximity to the electrodes, which should increase noise-immunity considerably. Such integration, which could include custom-fitting the electrode mounts for individual operators, will also minimize artifacts caused by even slight displacements of an electrode relative to the skin. Fortunately, some of the power supplies in fielded operational systems oscillate at frequencies considerably higher than the physiological signals of interest, so bandpass filters attentuate such noise sources more readily than the 60 Hz interference which can be a problem in the laboratory. Appropriate notch filters, akin to the 60 Hz filters used in many conventional amplifiers, can also be custom-designed for specific operational settings, as long as the frequencies being attenuated are sufficiently disparate from the physiological spectrum of interest.

Physiological artifacts from the operator himself can be more troublesome. As alluded to above, electrophysiological recordings of one physiological parameter can be contaminated by other physiological parameters with overlapping frequency components. For example, EEG and ERP recordings can be contaminated by eye blinks, heart beat, and muscle artifacts. Furthermore, excessive sweating can elicit skin potentials that interfere with the physiological measures of interest or can cause electrodes to be more easily dislodged. These problems dictate the need for innovative electrode designs, well-integrated into the operators clothing and other equipment, as well as the need for "intelligent" digital filtering algorithms (e.g., Ref. 47) to rid the recording of artifact.

Real-time Turnaround. As discussed in the "Areas of Application" section, many potential uses of physiological measures in operational settings do not require real-time turnaround of data analyses. In fact, most recordings to date in simulators or fielded systems have stored the amplified physiological signs on either analog or digital media for off-line analysis. Only recently have systems appeared with some on-board computing power (e.g. the SSPIDR), but even here the decision-making capability has thus far been limited to making intelligent decisions about when to store data into the limited-capacity memory for off-line analysis. The possibilities for real-time analysis of physiological data, the use of derived measures for real-time decision-making, and the realization of closed-loop feedback based on the resultant decisions as an input to adaptive systems are areas that need to be pursued more aggressively. A reasonable way to proceed on this front would seem to be an initial focus on the development of pattern recognition algorithms for "single trial" extraction of useful indices from records of ongoing physiological activity, followed by non-real-time demonstrations of how these derived indices would be used for making useful decisions about operator status. Only then need there be an attempt to "speed-up" this process to real-time, perhaps by implementing the mature algorithms in special-purpose hardware.

Knowledge-based Interpretation of Physiological Measures. Whether or not real-time turnaround is required for certain applications in operational settings, there will certainly be the need for more automated means of interpreting physiological data than are presently available. Expert system techniques for encoding knowledge and applying decision rules offer possibilities as a framework for such automated interpretation, although it is not yet clear how complicated the decision-rules and contingencies will need to be. It is apparent, given the aforementioned cautions that have been raised about inferring mental states from physiological measures alone (Ref. 2), that it will be necessary to take into account simultaneously derived measures of operator behavior and system performance as a whole. Very little work has been done in modeling the integration of physiological, behavioral and system performance data. The paper by Samaras in the present Proceedings offers one possible framework for such an integration. Appropriate decision rules relating changes in physiological signs to mental states or predicted performance can be derived initially from the biomedical and psychophysiological literatures. However, refinements of these decision rules and proof-of-concept demonstrations will likely require the use of realistic scenarios in simulator environments.

SUMMARY OF AREAS FOR FURTHER DEVELOPMENT

The foregoing discussion has attempted to provide an overview of the state-of-the-art and the challenges that lie ahead "in the field," as physiological monitoring technology expands from the laboratory into operational settings. Although valid physiological measures have been recorded already in a number of demanding operational settings, including advanced cockpits, the methodologies for implementing such measures have been largely special-purpose and cumbersome. The successes to-date merely foreshadow the possibilities that exist, as conceptual and engineering advances continue. The following list summarizes a number of the areas that are fertile ground for further development:

- Advances in physiological sensor design and better ways of mounting electrodes in an operator's helmet, clothing, or other gear.

- Further miniaturization of amplifiers, digitizers, storage media, and telemetry equipment, along with design features to maximize noise-immunity and integration into the operator's physical environment.

- Digital filtering algorithms to minimize the contamination of recordings by artifacts, both those due to electrical sources in the environment and those due to physiological sources within the subject.

- Special purpose data analysis software or firmware that can process the recorded signals in near real-time.

- Modeling of mental states and task environments to allow physiological measures to be taken into account, along with behavioral, subjective, and system performance measures, in interpreting and predicting performance.

- Empirical work to develop decision algorithms for inferring the operational significance of operator physiological changes and for "closing the loop" between man and machine.

REFERENCES

1. Coles, M. G. H., Donchin, E. & Porges, S. W. (Eds.): Psychophysiology: Systems, Processes, and Applications. New York: Guilford, 1986.

2. Johnson, L. C.: Use of Physiological Measures to Monitor Operator State. In F. E. Gomer (Ed.), Biocybernetic Applications for Military Systems. Report No. MDC E2191, St. Louis: McDonnell Douglas Astronautics Company, 1980, pp. 101-131.

3. Zacharias, G. L.: Physiological Correlates of Mental Workload. Report No. 4308, Cambridge: Bolt, Beranek & Newman, Inc., 1980.

4. O'Donnell, R. D.: Contributions of Psychophysiological Techniques to Aircraft Design and Other Operational Problems. NATO AGARDograph No. 244, 1979.

5. Gomer, F. E. (Ed.): Biocybernetic Applications for Military Systems. Report No. MDC E2191, St. Louis: McDonnell Douglas Astronautics Company, 1980.

6. Gevins, A. S., Doyle, J. C., Cutillo, B. A., Schaffer, R. E., Tannehill, R. S., Ghannam, J. H., Gilcrease, V. A., & Yeager, C. L.: Electrical Potentials in Human Brain During Cognition: New Method Reveals Dynamic Patterns of Correlation. Science, vol. 213, 1981, pp. 918-922.

7. Lindsley, D. B.: Psychological Phenomena and the Electroencephalogram. Electroencephalography and Clinical Neurophysiology, vol. 4, 1952, pp. 443-456.

8. Malmo, R. B.: Activation: A Neurophysiological Dimension. Psychological Review, vol. 66, 1959, pp. 367-386.

9. Duffy, E.: Activation. In N. S. Greenfield & R. A. Steinbach (Eds.), Handbook of Psychophysiology. New York: Holt, Rinehart and Winston, 1972, pp. 577-622.

10. Sem-Jacobsen, C. W.: Electroencephalographic Study of Pilot Stresses in Flight. Aerospace Medicine, vol. 30, 1959, pp. 797-801.

11. Sem-Jacobsen, C. W., & Sem-Jacobsen, E. E.: Selection and Evaluation of Pilots for High Performance Aircraft and Spacecraft by Inflight EEG Study of Stress Tolerance. Aerospace Medicine, vol. 34, 1963, pp. 605-609.

12. Donchin, E., Ritter, W., & McCallum, W. C.: Cognitive Psychophysiology: The Endogenous Components of the ERP. In E. Callaway, P. Tueting, & S. H. Coslow (Eds.), Event-Related Brain Potentials in Man. New York: Academic Press, 1978, pp. 349-411.

13. Hillyard, S. A. & Hansen, J. C.: Attention: Electrophysiological Approaches. In M. G. H. Coles, E. Donchin, & S. W. Porges (Eds.), Psychophysiology: Systems, Processes and Applications. New York: Guilford Press, 1986, pp. 227-243.

14. Sutton, S. Braren, M., Zubin, J. & John, E. R.: Evoked Potential Correlates of Stimulus Uncertainty. Science, vol. 150, 1966, pp. 1187-1188.

15. Ritter, W., Simson, R., & Vaughan, H. G., Jr.: Association Cortex Potentials and Reaction Time in Auditory Discriminations. Electroencephalography and Clinical Neurophysiology, vol. 33, 1972, pp. 547-557.

16. Walter, W. G., Cooper, R., Aldridge, V. J., McCallum, W. C. & Winter, A. L.: Contingent Negative Variation: An Electrical Sign of Sensori-motor Association and Expectancy in the Human Brain. Nature, vol. 203, 1964, pp. 380-384.

17. McCarthy, G. & Donchin E.: A Metric for Thought: A Comparison of P300 Latency and Reaction Times. Science, vol. 211, 1981, pp. 77-79.

18. Karis, D., Chesney, G. L., & Donchin, E.: "...twas ten to one; And yet we ventured...": P300 and Decision Making. Psychophysiology, vol. 20, 1983, pp. 260-268.

19. Wilson, G. F. & O'Donnell, R. D.: Steady-State Evoked Responses: Correlations with Human Cognition. Psychophysiology, vol. 23, 1986, pp. 57-61.

20. Beideman, L. R. & Stern, J. A.: Aspects of the Eyeblink During Simulated Driving as a Function of Alcohol. Human Factors, vol. 19, 1977, pp. 73-77.

21. Bauer, L. O., Strock, B. D., Goldstein, R., Stern, J. A. & Walrath, L. C.: Auditory Discrimination and the Eyeblink. Psychophysiology, vol. 22, 1984, pp. 636-641.

22. Goldstein, R., Walrath, L. C., Stern, J. A., and Strock, B. D.: Blink Activity in a Discrimination Task as a Function of Stimulus Modality and Schedule Presentation. Psychophysiology, vol. 22, 1985, pp. 629-635.

23. Bauer, L. O., Goldstein, R., & Stern, J. A.: Effects of Information-Processing Demands on Physiological Response Patterns. Human Factors, vol. 29, 1987, pp. 213-234.

24. Harris, R. L., Sr., Tole, J. R., Stephens, A. T., & Ephrath, A. R.: Visual Scanning Behavior and Pilot Workload. Aviation, Space and Environmental Medicine, vol. 53, 1982, pp. 1067-1072.

25. Beatty, J.: Phasic Not Tonic Pupillary Responses Vary with Auditory Vigilance Performance. Psychophysiology, vol. 19., 1982, pp. 167-172.

26. Casali, J. G., & Wierwille, W. W.: A Comparison of Rating Scale, Secondary-Task, Physiological, and Primary-Task Workload Estimation Techniques in a Simulated Flight Task Emphasizing Communications Load. Human Factors, vol. 25, 1983, pp. 623-641.

27. Robinson, E. R. N.: Biotechnology Predictors of Physical Security Personnel Performance: I. A Review of the Stress Literature Related to Performance. Report No. NPRDC TN-83-9, Navy Personnel Research and Development Center, San Diego, Calif., 1983.

28. Lacey, J. I.: Somatic Response Patterning and Stress: Some Revisions of Activation Theory. In M. H. Appley & R. Trumbull (Eds.), Psychological Stress. New York: Appleton-Century-Crofts, 1967.

29. Roscoe, A. H.: Heart-Rate as an In-Flight Measure of Pilot Workload. In M. L. Frazier & R. B. Crombie, (Eds.), Proceedings of the Workshop on Flight Testing to Identify Pilot Workload and Pilot Dynamics. Edwards Air Force Base: AFTEC-TR-82-5, 1982, pp. 338-349.

30. Hart, S. G., Hauser, J. R., & Lester, P. T.: Inflight Evaluation of Four Measures of Pilot Workload. Proceedings of the Human Factors Society -- 28th Annual Meeting, 1984, pp. 945-949.

31. Sayers, B. McA.: Physiological Consequences of Informational Load and Overload. In P.H. Venables & M.J. Christie (Eds.), Research in Psychophysiology. New York: Wiley, 1975, pp. 95-124.

32. Veldman, J. B. P., Mulder, L. J. M., Mulder, G., & van der Heide, D.: Attention, Effort and Sinus Arrhythmia: How Far Are We? In J. F. Orlebeke, G. Mulder, and L. J. van Doornen (Eds.), Psychophysiology of Cardiovascular Control. New York: Plenum Press, 1985, pp. 407-424.

33. Mulder, G.: Sinusarrhythmia and Mental Workload. In N. Moray (Ed.), Mental Workload: Its Theory and Measurement, New York: Plenum Press, 1979, pp. 327-344.

34. Porges, S. W.: Respiratory Sinus Arrhythmia: An Index of Vagal Tone. In J. F. Orlebeke, G. Mulder, and L. J. van Doornen (Eds.), Psychophysiology of Cardiovascular Control. New York: Plenum Press, 1985, pp. 437-450.

35. Hatch, J. P., Klatt, K., Porges, S. W., Schroeder-Jasheway, L. & Supik, J. D. The Relation Between Rhythmic Cardiovascular Variability and Reactivity to Orthostatic, Cognitive, and Cold Pressor Stress. Psychophysiology, vol. 23, 1986, pp. 48-56.

36. Burton, R. R.: Human Responses to Repeated High-G Simulated Aerial Combat Maneuvers. Aviation, Space, and Environmental Medicine, vol. 51, 1982, pp. 1185-1192.

37. Poppen, J. R. & Drinker, C. K.: Physiologic Effects and Possible Methods of Reducing Symptoms Produced by Rapid Changes in Speed and Direction of Airplane as Measured in Actual Flight. Applied Physiology, vol. 215, 1950.

38. Roman, J., Older, H., & Jones, W. L.: Flight Research Program: VII. Medical Monitoring of Navy Carrier Pilots in Combat. Aerospace Medicine, vol. 38, 1967, pp. 133-139.

39. Rugh, J. D., Wichman, H., & Faustman, W. O.: Inexpensive Technique to Record Respiration During Flight. Aviation, Space and Environmental Medicine, vol. 48, 1977, pp. 169-171.

40. Williges, R. C., & Wierwille, W. W.: Behavioral Measures of Aircrew Mental Workload. Human Factors, vol. 21, 1979, pp. 549-574.

41. Eason, R. G., Beardshall, A. & Jaffe, S.: Performance and Physiological Indicants of Activation in a Vigilance Situation. Perceptual and Motor Skills, vol. 20, 1965, pp. 3-13.

42. Corkindale, K. G., Cumming, F. G., & Hammerton-Fraser, A. M.: Physiological Assessment of Pilot Stress During Landing. In Measurement of Aircrew Performance, AGARD Conference Proceedings, CP#56, Brooks Air Force Base, Texas, 1983.

43. Jex, H. R. & Allen, R. W.: Research on a New Human Dynamic Response Test Battery; Part II. Psychophysiological Correlates. Proceedings of the Sixth Annual Conference on Manual Control, Wright-Patterson Air Force Base, Ohio: Air Force Institute of Technology, 1970, pp. 743-777.

44. Okada, Y. C., Kaufman, L., & Williamson, S. J.: The Hippocampal Formation as a Source of the Slow Endogenous Potentials. Electroencephalography and Clincal Neurophysiology, vol. 55, 1983, pp. 417-426.

45. Hancock, P. A.: Task Categorization and the Limits of Human Performance in Extreme Heat. Aviation, Space, and Environmental Medicine, vol. 53, 1982, pp. 778-784.

46. Call, D. W., Kelly, D. M., & Robertson, D. G.: A Self-Contained, Man-Borne Biomedical Instrumentation System in the Flight Testing of Naval Weapons Systems. In M. L. Frazier & R. B. Crombie, (Eds.), Proceedings of the Workshop on Flight Testing to Identify Pilot Workload and Pilot Dynamics, Edwards Air Force Base: AFTEC-TR-82-5, 1982, pp. 318-321.

47. Widrow, B., Glover, J. R., McCool, J. M., Kaunitz, J., Williams, C. S., Hearn, R. H., Zeidler, J. R., Dong, E., & Goodlin, R. C.: Adaptive Noise Cancelling: Principles and Applications. Proceedings of the IEEE, vol. 63, 1975, pp. 1692-1716.