Text of remarks made at the 1982 American Psychological Association Annual Meeting. This material will appear as an Air Force Aerospace Medical Research Laboratory Technical Report.

A CONCEPTUAL FRAMEWORK FOR DEVELOPMENT OF A WORKLOAD ASSESSMENT METHODOLOGY

F. Thomas Eggemeier
Wright State University
and
Systems Research Laboratories, Inc.

and

Robert D. O'Donnell
USAF Aerospace Medical Research Laboratory

1982

ABSTRACT

Based on a review of the current literature, a conceptual framework which incorporates major elements related to operator workload has been developed. The framework treats workload as a multidimensional construct with important physiological, subjective, and behavioral components. An important implication of the framework is that, at present, a comprehensive workload assessment methodology should include a number of measures, including subjective, physiological, and performance-based metrics.

INTRODUCTION

A primary concern of human factors engineering during system development and evaluation is to assure that the performance demands imposed by a

APPENDIX D

REFERENCES

system do not exceed the human operator's capacity to process information. Mental workoad is the term which has been used in referring to the degree or percentage of the operator's information processing capacity which is expended in meeting system demands. The growing complexity of modern systems and associated increases in workload have increased the likelihood of exceeding or approaching the limitations of an operator's processing capacity. As a consequence, the need for reliable and sensitive methods of assessing the workload imposed by alternative design options has assumed increased importance in recent years.

In order to address the need for viable workload metrics, a large number of individual measurement techniques have been developed and documented in a number of reviews (e.g., Gartner and Murphy, 1976; Williges and Wierwille, 1979; Moray, 1982). Williges and Wierwille (1979), for example, identified 28 specific workload assessment techniques which have been proposed in the recent literature. Despite the fact that there are numerous individual techniques, most workload measures can generally be classified as belonging to one of the three major categories:

1. Subjective measures, such as rating scales, which require the operator to rate or somehow characterize the subjective workload associated with performance of a particular task or with a system design option.

2. Psychophysiological measures, such as heart rate variability, which derive an index of workload from some aspect of the operator's physiological response to task or system demands.

3. Performance-based measures, such as secondary task methodology, which use some aspect of the operator's capability to perform within the system as a measure of workload.

All three classes of measures have been used extensively during the past 10 years with varying degrees of success. Given the large number of measures, a major issue in developing a workload assessment methodology is choosing those measures that should be incorporated into the methodology.

Unfortunately, although there is considerable agreement in the literature regarding the importance of workload, there has not been substantial agreement concerning the most appropriate means by which to assess it. Much of the failure to agree on the appropriate assessment technique stems from the fact that workload is a multidimensional concept (e.g., Johannsen, Moray, Pew, Rasmussen, Sanders, and Wickens, 1979; Sanders, 1979; Williges and Wierwille, 1979; Sheridan and Simpson, 1979; White, 1971) which has been used in referring to several different aspects of system or task demand, operator effort and information processing capacities, operator performance, and systems performance (e.g., Sheridan and Stassen, 1979).

In order to provide a means for organizing the numerous factors associated with workload, a general conceptual framework which addresses major elements of the workload construct has been developed. The major objective in assembling the framework was to help guide the development of a comprehensive workload assessment methodology which would tap essential elements of the workload construct. The framework is largely based on current conceptualizations of workload and theories regarding the nature of capacity limitations within the human information processing system. The purpose of this paper is to provide an overview of the conceptual framework and to discuss implications of the framework for workload assessment.

CONCEPTUAL FRAMEWORK FOR WORKLOAD

Despite the diversity in emphasis among various approaches to workload, three general elements which represent major factors in most current theoretical treatments can be identified. These general elements include:

1. Some characterization of the demands placed on the operator by the system or task.

2. Some expression of the capacity or effort expenditure required by the operator to deal with the demands.

3. The level of operator performance that results from the interaction of task demands and capacity/effort expenditure.

Figure 1 illustrates these major elements and their interrelationship at a very general level.

Insert Figure 1 here

As indicated in Figure 1, system demands are imposed on the operator and must be related to or mapped onto the operator's processing capacity or resources. Depending upon the efficiency of the operator in mapping demands onto available capacities and the degree of effort expended by the operator, a particular level of performance results. Effort is intended to represent a construct similar to that described by Kahneman (1973) and Jahns (1973), and represents a nonspecific input which is required to activate information processing structures. Most conceptualizations of workload include each of these elements, and a variety of workload assessment procedures have been developed to evaluate aspects of each element.

Several current models and theoretical statements concerning workload (e.g., Jahns, 1973; Welford, 1978; Sanders, 1979; Sheridan and Simpson, 1979; Johannsen et al., 1979) treat major elements outlined in Figure 1 in more detail. These models and statements generally maintain that workload is clearly a multidimensional construct that reflects the interaction of such elements as task and system demands, operator processing capacities and effort, subjective performance criteria, operator information processing strategies, and operator training or experience.

In addition to suggesting that workload itself is multidimensional, a number of investigators have also maintained that several elements of workload depicted in Figure 1 are themselves multidimensional.

Jahns (1973), for instance, categorized sources of input load or system demand into three classes: (1) environmental, (2) situational, and (3) procedural. Environmental demands included factors such as temperature, humidity, noise, and acceleration which can serve as sources of load for an operator. Demands that were characterized as situational were composed of elements such as display and control characteristics, display-control arrangement, vehicle dynamics, and crewstation volume. These types of demand reflect traditional human factors engineering concerns, and play a central role in determining the overall level of input load experienced by an operator. Finally, procedural demands included elements such as mission or task duration, standard system operating procedures, the mission itself, and briefings/instructions given to the operator. Jahns does not specify how the various sources of demand combine to determine overall levels of load, but it is clear that input load is considered to be multidimensional and is described as a vector rather than a scalar quantity.

A similar conclusion has been reached by Johannsen et al. (1979) with respect to the concept of effort (Kahneman, 1973; Jahns, 1973). Johannsen et al. (1979) pointed out that the notion of effort is extremely complex and can be interpreted in several different ways. One sense of effort identified by Johannsen et al. is related to the physiological activation as measured by a number of indices (e.g., muscle tension, respiration rate) that occurs when an operator is exposed to progressive increases in mental load. A second related sense of effort is the subjective feeling experienced by an operator under a high load condition. Johannsen et al. assume that such feelings represent the products of muscular tension and changes in physiological variables such as blood pressure and heart rate. It is noted that an operator may actually feel loaded and effortful despite the fact that there is no change in the adequacy of performance. This can be explained by assuming that increases in task difficulty lead to a period during which the operator is working harder in order to preclude performance decrements. The clear implication is that effort has both physiological and subjective dimensions, each of which may reflect increases in load prior to any actual decrements in operator performance.

Operator processing capacity which is depicted in Figure 1 has also been characterized as multidimensional in some recent descriptions of the information processing system. This position assumes that the information processing system may be described as a series of internal processing structures, each with its own processing capacity or resources which are not exchangeable with any other structure. Current theoretical positions that are consistent with this model are the multiple resources or structure-specific resource models of processing capacity (e.g., Wickens, 1979, 1981; Navon and Gopher, 1979; Sanders, 1979; North, 1977; Kantowitz and Knight, 1976). According to this theory, a considerable amount of capacity may remain unused in responding to a particular task's demands, because ony a limited number of processing structures may be involved in that response. An important consequence of this position identified by Sanders (1979), Wickens (1979), and Gopher (1978), is that mental load cannot be conceptualized as a single dimension; and ultimately, a task may have to be described in terms of multidimensional patterns of mental load. A central question in workload specification becomes one of determining the extent to which various processing resources are involved in a task which, in turn, specifies the pattern of mental load. Several theorists (e.g., North, 1977; Sanders, 1979) have suggested candidate dimensions of resources that could be used in describing the resource demand composition of a task. The most comprehensive position, however, is that of Wickens (1981) who, on the basis of available evidence, has identified three primary dichotomous dimensions that appear to define separate resources. These dimensions include:

1. Stages of information processing (perceptual/central processing operations versus response selection and execution).

2. Modalities of perception (auditory versus visual).

3. Codes of information processing and response (spatial-manual versus verbal-vocal).

Under the system proposed by Wickens, an adequate description of the load imposed by a task should include specification of its demand composition on each of the three dimensions. The implication of the multiple resources theory is clear: the capacity to process information is multidimensional and characterizations of the load imposed by a task should reflect that multidimensionality.

The general conceptual framework outlined previously can now be modified to be more descriptive of elements related to workload. Figure 2 illustrates the modified conceptual framework. In the modified framework, workload and operator performance continue to represent the product of an interaction of several factors. In accordance with Jahns' position, system demands are represented as multidimensional, including the classes of environmental, situational, and procedural load. Effort is also conceptualized as reflecting several dimensions, including strong physiological and subjective components. Likewise, operator processing resources have been revised to reflect several of the dimensions represented in the multiple resources view of the information processing system.

Insert Figure 2 here

The major conclusion that follows from the present framework is that workload, as currently conceptualized, represents a multidimensional construct that includes important physiological, subjective, and behavioral components. The multidimensionality is reflected not only in the interaction of several elements to determine levels of load, but also in the multidimensionality of several elements themselves. The multidimensional nature of the framework has important implications for workload assessment, and these are discussed in the next section.

IMPLICATIONS OF THE CONCEPTUAL FRAMEWORK FOR WORKLOAD ASSESSMENT

A major implication of the current framework is that no single measurement technique will provide a comprehensive means for assessment of load. Since current theory generally maintains that there are important physiological, subjective, and behavioral components of load, a comprehensive approach to assessment of workload should include physiological, subjective, and behavioral or performance-based measures. At present, it is not clear how such measures might ultimately be combined to provide a multidimensional index of load, but it should be clear from the foregoing discussion that subjective and physiological measures can potentially provide information not afforded by performance-based measures, and vice versa. Therefore, it appears that the most viable approach to comprehensive workload assessment would be a battery of measures, including subjective, physiological, and performance-based components which could be applied to derive indices of several components of load. Similar conclusions have been drawn recently by several others. Johannsen et al. (1979), for example, concluded that workload contains behavioral, performance, physiological, and subjective components and indicated that appropriate measures would be required for each component. Williges and Wierwille (1979), in their review of behavioral workload assessment procedures, maintain that due to the multidimensionality of workload, it appears unlikely that any single measure will be completely sufficient for characterizing load. Williges and Wierwille conclude that multiple measures, includng dimensions of subjective opinion, spare mental capacity, primary tasks, and physiological correlates need to be considered. Similarly, Jahns (1973) noted that because of the complex interactions involved in determining levels of operator effort, a broad spectrum of measurement techniques for operator workload need to be investigated. White (1971) also indicated that workload is multidimensional and, at present, cannot be defined adequately in terms of any single measure.

Further support for the necessity of considering multiple measures of load comes from the fact that when several measures of workload are applied in a situation, they commonly exhibit some degree of dissociation (e.g., Borg, 1978; Hicks and Wierwille, 1979; Dornic and Andersson, 1980; Wickens and Derrick, 1981; and Moray, 1982). This type of result can, of course, be interpreted within the multidimensional workload framework outlined above, since different measurement techniques can be assumed to be maximally sensitive to different dimensions of load. Wickens and Derrick (1981), for instance, have noted that a lack of correspondence among workload measures can be attributed to the multidimensional nature of information processing resources that are assumed to underlie performance and workload. Based on the results of several experiments in which dissociation occurred between physiological, primary task, subjective, and secondary task measures, Wickens and Derrick proposed that some measures may be considered more generally sensitive to overall levels of demand for resources anywhere within the information processing system, while others are more diagnostic in the sense of specifically reflecting demands imposed on particular resources (e.g., percpetual versus motor output) within the processing system. It was suggested that some physiological measures such as heart rate variability might be sensitive to the total demand placed on all resources of the system, even if the demand was imposed such that no single resource was overloaded and no performance decrement occurred. Subjectve measures were also viewed as generally sensitive to demands imposed anywhere in the system, while secondary task measures were thought to be more diagnostic in providing information regarding the specific resources demanded by a task.

Wickens and Derrick concluded that since individual measures might provide different types of information about load, choice of measure might be dictated in part by the purpose of taking the measure. If the total load imposed by a task or system were to be assessed, some physiological or subjective techniques might be used in conjunction with primary task performance measures. However, if the specific locus of an overload (e.g., central processing versus response execution) were to be identified in order to provide more diagnostic information to a human factors engineer, then secondary task methodology might be the more appropriate choice. Note that in this type of scheme, several measurement techniques could be used in a complementary fashion, with some subjective or physiological measures providing an initial index of overall task or system workload, and secondary tasks or other physiological measures (e.g., evoked cortical potential) being used subsequently to develop more precise information regarding the locus of specific overloads.

The clear implication that follows from the conceptual framework and from the noted dissociation of workload measures is that, at present, a comprehensive workload assessment methodology will require physiological, subjective, and performance-based measurement techniques. The major goals for research in the workload metric development area suggested by the framework are to: [1] initially identify the most sensitive measure(s) within each category of assessment technique; and [2] conduct a systematic comparison of the information provided by each category of assessment technique so that optimal combination(s) of measures required for a comprehensive workload assessment methodology can ultimately be established.

REFERENCES

Borg, G. Subjective aspects of physical and mental load. Ergonomics, 1978, 21, 215-220.

Dornic, S. and Andersson, O. Difficulty and effort: A perceptual approach. Stockholm, Sweden: Reports from the Department of Psychology, the University of Stockholm, Report No. 566, November 1980.

Gartner, W. B. and Murphy, M. R. Pilot workload and fatigue: A critical survey of concepts and assessment techniques. Washington, D.C.: National Aeronautics and Space Administration, Report No. NASA-TN-D-8365, November 1976.

Hicks, T. G. and Wierwille, W. W. Comparison of five mental workload assessment procedures in a moving-base driving simulator. Human Factors, 1979, 21, 129-143.

Jahns, D. W. A concept of operator workload in manual vehicle operations. Meckenheim, Germany: Forschungsinstitut fur Anthropotechnik, Report No. 14, 1973.

Johannsen, G., Moray, N., Pew, R., Rasmussen, J., Sanders, A., and Wickens, C. Final report of the experimental psychology group. In: N. Moray (ed.), Mental Workload: Its theory and measurement. New York: Plenum Press, 1979.

Kahneman, D. Attention and effort. Englewood Cliffs, New Jersey: Prentice-Hall, 1973.

Kantowitz, B. H. and Knight, J. L. Testing tapping timesharing, II: Auditory secondary task. Acta Psychologica, 1976, 40, 342-362.

Moray, N. Subjective mental workload. Human Factors, 1982, 24, 25-40.

Navon, D. and Gopher, D. On the economy of the human processing system, Psychological Review, 1979, 86(3), 214-255.

North, R. A. Task components and task demands as factors in dual-task performance. University of Illinois, Aviation Research Laboratory, Report No. ARL-77-2, January 1977.

Sanders, A. F. Some remarks on mental load. In: N. Moray (ed.), Mental workload: Its theory and measurement. New York: Plenum Press, 1979.

Sheridan, T. B. and Simpson, R. W. Toward the definition and measurement of the mental workload of transport pilots. Cambridge, Massachusetts: Massachusetts Institute of Technology Flight Transportation Laboratory Report, FTL Report R79-4, January 1979.

Sheridan, T. B. and Stassen, H. G. Definitions, models and measures of human workload. In: N. Moray (ed.), Mental workload: Its theory and measurement, New York: Plenum Press, 1979.

Welford, A. T. Mental workload as a function of demand, capacity, strategy, and skill. Ergonomics, 1978, 21, 151-167.

White, R. T. Task analysis methods: Review and development of techniques for analyzing mental workoad in multiple task situations. St. Louis, Missouri: McDonnell-Douglas Corporation, MDC J 5291, September 1971.

Wickens, C. D. Measures of workload, stress, and secondary tasks. In N. Moray (ed.), Mental workload: Its theory and measurement. New York: Plenum Press, 1979.

Wickens, C. D. Processing resources in attention, dual task performance, and workload assessment. Engineering Psychology Research Laboratory Technical Report EPL-81-3, University of Illinois, 1981.

Wickens, C. D. and Derrick, W. Workload measurement and multiple resources. In: Proceedings of the 1981 IEEE Conference on Cybernetics and Society, 1981, 600-603.

Williges, R. C. and Wierwille, W. W. Behavioral measures of aircrew mental workload. Human Factors, 1979, 21, 549-574.