

Practical Guidelines For Workload Assessment

Andrew J. Tattersall

University of Wales, Cardiff

Introduction

There are numerous factors that need to be taken into account in a comprehensive human factors certification process. This paper deals with one major factor which is a central area for human factors concern when humans are required to interact with complex systems, namely workload. If the process of human factors certification of systems is to succeed, then workload assessment must be incorporated into the evaluation and certification process. Proper and effective evaluation will ensure that the workload experienced by users of any system, in aviation or otherwise, is taken account of in system design and development.

There is now a vast amount of literature on workload assessment. This interest has been stimulated primarily because of the need to design complex task environments that do not place undue demands and requirements on the human operator. The principal applications have typically been in process control and aviation settings such as air traffic control and aircraft cockpit design. This paper will address a number of practical issues that need to be taken into account when any evaluation of existing systems or future systems is carried out.

There is little dispute that workload is a multidimensional concept (Damos, 1991; Gopher & Donchin, 1986; Moray, 1982; O'Donnell & Eggemeier, 1986), but one distinction that is not very often made explicit is between acute and chronic dimensions of the impact of workload. If workload is defined in terms of the costs that operators incur in performing tasks (Kramer, 1991), then the distinction is even more apposite. An understanding of this distinction will, to a certain extent, aid evaluations of the nature of workload in work settings.

First, if one is concerned with the acute effects of workload, then the main focus of the research will be on the interference between tasks in dual- or multiple-task situations. A principal question to be asked is whether the tasks are too demanding in terms of the human information processing requirements, in which case performance on one or more of the primary work tasks may be degraded. The logic of many of the approaches in this area is based upon multiple-resource theory (e.g. Wickens, 1984, 1991), in which it is proposed that there are a variety of processing resources that are limited in their capacity. The extent to which tasks will interfere with each other when carried out concurrently will depend upon the extent to which they compete for common resources. Evaluation of workload in this case may include primary and secondary task performance measures, subjective measures and certain psychophysiological measures. The ultimate intention of such approaches is to predict performance in multiple-task situations.

Second, if one is interested in chronic symptoms of heavy workload then the main concerns are with the effects of managing the demands of work over a day, a week or a prolonged period of time. The after-effects of work are an important consideration. In other words, to what extent is the physiological and emotional state of an individual affected over a period of time, and does subsequent performance at work show any decrement due to these changes in state? For example, individuals may become increasingly fatigued because of the sustained demands in the job that need to be managed effectively. This may result in breakdown of skills over the longer-term, and current models of workload are unable to predict these outcomes with any degree of certainty. Typical assessment techniques might again involve the use of performance measures, subjective measures and physiological measures of workload. In addition, subjective and physiological measures of individual state are likely to be useful as there may be implications of the longer-term effects of workload for well-being, both emotional and physiological, health, performance and safety at work. Cross-sectional studies of occupational stress and health provide a useful way to gain background information about possible sources of stress, and the coping strategies and behavioural styles which might be generally effective in moderating the effects of work demands (e.g. Farmer, Belyavin, Berry, Tattersall & Hockey, 1990). However, it is also important to investigate the effects of exposure to different job demands in the longer-term. The development of a model of stress and workload based on observations of individual patterns of response to various demands will lead to the more accurate prediction of states or situations in which a breakdown in skills might occur.

Questions That Should Be Asked

Before any evaluation is carried out, a number of questions need to be addressed that will enable the appropriate measures to be taken for the particular problem that is to be tackled. An approach that has been used to great effect in assessing the usability of human-computer systems (Ravden & Johnson, 1989) is to provide users or designers of systems with a checklist of items to consider. The questions below are based on part of that checklist, and they may be useful for workload assessors to identify the key areas of concern about workload in the initial stages of an evaluation.

1. What is the general question? Is the concern with the overall system or a specific piece of equipment, and is the focus primarily on performance or on the health and safety of the workforce?
 - Is the primary concern with task design (the scheduling or allocation of tasks within or between individual operators' jobs)?
 - Is the primary concern with equipment design (for example, to evaluate the effects of the introduction and integration of a new item of technology, or to evaluate the difficulties of working with one item of equipment)?

- Is the primary concern with health and safety (the outcomes of working within an existing or proposed system perhaps as a result of sustained task demands, underload or boredom, etc.)?
2. Is there a specific problem? That is, do operators complain about specific tasks or functions?
- What are the best aspects of the system?
 - What are the worst aspects of the system?
 - What parts of the system give the most difficulty when carrying out the task?
 - Are there parts of the system that are confusing or difficult to understand?
 - What are the most common mistakes made in using the system?
 - What performance problems exist with the system?
 - What changes do operators suggest might be made to the system to make it more effective and usable?
3. What is the aim of measuring workload?

These questions will help to define the problem and the goals to be set for the workload assessment exercise. By focusing on these issues, the task of choosing a set of workload assessment techniques and specifying the research environment and design should be more straightforward and the results of studies easier to interpret and act upon.

Factors To Consider In Choosing A Particular Technique

Many different measures of workload have been developed, but their effective use in particular situations will depend on various factors, including their sensitivity to changes in demand, their ability to distinguish different kinds of demand, and their suitability or relevance to that situation. Extensive reviews of these techniques have been produced before, many of which discuss criteria for application (Gopher & Donchin, 1986; O'Donnell & Eggemeier, 1986; Hancock & Meshkati, 1988). The main factors to consider are as follows.

Validity. Moray (1988) has argued in a recent review of the development of mental workload research that, because no clear, precise definition of workload exists, it is difficult to establish the validity of different techniques. He suggests that the reliability of measures has to be sufficient for practical purposes until such a definition is agreed upon.

Reliability. Measures should be accurate and provide similar values from different operators doing the same task. There should be a good correlation between the values produced by different techniques if they are used to assess the same dimensions of workload. They should also have test-retest reliability, although as yet few reliability studies have been carried out.

Sensitivity. The concern here is with the effectiveness of the technique to discriminate between different levels of primary task load. In other words, does the technique actually measure changes in task demands and identify conditions of extreme workload?

Diagnosticity. Diagnosticity refers to the extent to which the technique is able to distinguish between different types of task demands and to identify the particular components within complex tasks that result in difficulty. Some techniques provide a general measure of resource allocation or effort, whereas others, such as secondary-task methodology, may be more sensitive to variations in different domains of processing. They may, for example, distinguish between verbal processing requirements and spatial processing requirements of tasks. Primary task measures of performance provide a global measure and are not really suitable for this purpose. Subjective measures, unless used to assess different task components, systems or functions for example, are generally not terribly diagnostic. The need for diagnosticity really depends upon the aims of the study. If the aim is to assess the introduction of a new piece of equipment or change in working procedure, then diagnosticity may not be critically important. If, on the other hand, there is a need to assess the demands of different control actions (e.g. manual or spoken) in relation to different modes of information presentation (e.g. visual or auditory), then the diagnosticity of the technique will be an important factor.

Intrusiveness. This refers to the extent to which the workload assessment technique disrupts the performance of the primary work task. The disruption could result from the use of obtrusive equipment or the application of a technique, or, in the case of secondary-task methodology in particular, simply the requirement to perform a concurrent task (Ogden, Levine & Eisner, 1979). If safety is a major concern (in air traffic control, for example), then clearly workload assessment techniques that may degrade performance should not be used. Simulation exercises may provide useful data should techniques which may be intrusive also provide other useful attributes.

Generality, Acceptability and Applicability. On a pragmatic level, it may be useful to choose a technique that can be used in different situations, perhaps so that comparisons can be made between different conditions. Ease of use and special requirements that may restrict the application of a technique, such as the need for special, expensive equipment, can be important considerations, as can the extent to which operators are accepting of the particular technique.

Workload Assessment Techniques

The main concern here is with what is termed 'mental workload' rather than physical workload, for which there are reliable and established assessment techniques available (e.g., Rohmert, 1987). The concept of mental workload has proved to be more difficult to define and measure, which causes some concern when the focus is on the tasks of air traffic controllers and pilots as

these tasks primarily involve cognitive processes rather than place great physical demands on personnel. It is not easy to estimate the demands of these tasks and therefore to predict the consequences of different levels of demand. We not only need valid, reliable and sensitive measures of workload, but also good methods based in sound theory to analyse the cognitive activity that is required to perform these types of complex tasks.

A further important factor to consider is that different types of measurement technique may relate to different dimensions of workload and therefore may provide a different perspective of the particular demands of the task. Indeed, some subjective rating techniques are designed to assess a number of different dimensions, such as time pressure, frustration or anxiety, and mental effort. Certain physiological measures may be most sensitive to one particular dimension, for example, heart rate variability may be more likely to be associated with changes in effort, whereas mean heart rate may reflect changes in anxiety or physical effort.

Furthermore, there needs to be consistency in the way that terminology is applied, and finally, the operational procedures should be standardised as much as possible.

Self Report Measures

Subjective measures are relatively easy to employ, and asking workers to rate the levels of demand they experience and their state of well-being and health at least has face validity. It is intuitively attractive simply to ask controllers or other workers about the levels of demands they are experiencing and the impact of work demands at different levels. A number of subjective workload assessment techniques have been developed. Among the validated scales that are widely used in aviation settings are the NASA Task Load Index (TLX) (Hart & Staveland, 1988), and the Subjective Workload Assessment Technique (SWAT) (Reid & Nygren, 1988). They both assess perceived workload on a number of dimensions, usually after the task has been performed. Nygren (1991) has suggested that they are both useful measures of workload and that SWAT is sensitive at both individual and group levels. Hill, Iavecchia, Byers, Bittner, Zaklad & Christ (1992) compared SWAT, TLX, the Modified Cooper-Harper scale (Wierwille & Casali, 1983), and the Overall Workload scale (Vidulich & Tsang, 1987) for sensitivity, operator acceptance, response requirements, and any special procedures they require. All were found acceptable and sensitive to different levels of workload. Nygren (1991) points out, however, that the psychometric properties of these scales need to be fully understood, in addition to their implications for task performance, before they are applied extensively.

A recent development in this area has been the attempt to design subjective techniques which provide ratings of workload during primary task performance rather than after carrying out tasks. One such example is the Instantaneous Self Assessment technique (ISA), which was initially designed for use with air traffic control tasks (Hulbert, 1989; Jordan, 1992). Few evaluation studies have been carried out, but it appears to be a relatively sensitive measure of workload (Tattersall & Foord, 1993). A similar technique was used by Rehmann, Stein and Rosenberg (1983), who suggested that gaining concurrent workload evaluations was more accurate than post-task ratings. In complex tasks which involve multiple elements or phases, the ratings may be more clearly related to changing task demands than retrospective ratings. Tattersall and Foord (1993) found, however, that ISA responses, although correlated with other subjective workload measures, interfered to a certain extent with the primary tracking task, whether responses were made by speech or manually.

One problem with subjective measures has been their diagnosticity and to a certain extent their reliability, whereas their validity and sensitivity have been fairly well established. A

further worry is that they are not always found to correlate with measures of performance. Tasks that are performed better are sometimes found to have higher ratings of workload (e.g. Yeh & Wickens, 1988).

Self report scales of a different type have been developed to assess mood (e.g., Mackay, Cox, Burrows & Lazzerini, 1978; Warr, 1989), and longer-term health. Importantly, significant relationships have been found between subjective responses and specific physiological responses, such as that between cortisol and subjective distress (Frankenhaeuser and Johansson, 1986), and effort and heart rate variability (Aasman, Mulder & Mulder, 1987; Vicente, Thornton & Moray, 1987). If one is interested in the relationships between workload, performance and individual state, then repeated measurement of mood and other variables will be likely to be necessary.

Physiological Measures

These techniques include measures of cardiac function, brain function and other physiological processes. In terms of ECG measures, a number of studies now suggest that the power in the mid-frequency band of the heart rate variability (HRV) spectrum (0.07-0.14 Hz) is related to the level of mental effort invested in a task by an individual (Tattersall & Hockey, 1990; Aasman, Mulder & Mulder, 1987; Mulder, 1980; Vicente, Thornton & Moray, 1987). Such variability has been found to decrease as a function of task difficulty in a number of laboratory tasks such as tracking, memory search and classification tasks (Mulder, 1980). It is argued that bursts of suppressed vagal control correspond to periods of momentary effort or controlled processing. Mean heart rate may offer a more sensitive measure of response load, and is certainly influenced by physical activity and perhaps anxiety, which may limit its usefulness in relation to workload. The assessment of pupil dilation (Beatty, 1982) and EEG measures, including evoked potentials have also been used effectively, particularly in laboratory situations, but the advantage of measures of cardiac function are that they can be continuously and independently applied without intrusion to the primary task.

Other techniques, such as analyses of urine and blood provide measures of changes in physiological state through assessment of cortisol, adrenaline and noradrenaline excretion. Urine analysis allows measures of longer-term changes in state through assessment of cortisol and catecholamine concentration. Sustained stress states tend to show increased levels of these hormones. Blood or saliva samples may provide shorter-term measures of fluctuations in state.

There are potential problems with the diagnosticity of physiological measures (Wierwille & Casali, 1983), and a further problem to be aware of is that physiological processes are sensitive to the effects of physical activity and to emotional factors that may have an effect on physiological functions.

Performance

Two major approaches to performance assessment are primary task techniques and secondary task techniques. Primary-task measures are normally only useful for giving an indication of the impact of gross demands. These measures may be easy to obtain in some situations, but in others, such as air traffic control, it is difficult to generate a simple measure of a controller's level of performance that would meet the criteria outlined earlier. If one examined the safety record in air traffic control for example, one might conclude that workload is only a minor

problem. However, such a measure may not be immediately sensitive to the effects of changes in task load or working procedures, and will only give a crude indication of the cumulative effect of sustained and high task demands over a long period. Task strategies may differ between skilled and inexperienced operators, and the health and state of the operator may determine the perceived difficulty of a task. The effects of these factors may only be detected by primary task measures once performance suffers or errors are made. From a safety perspective this may be already too late a stage to investigate the adequacy of human factors aspects of the system.

Secondary task measures may be more sensitive to changes in demand or working procedures but unless the allocation of information processing resources to the two tasks is controlled it can be difficult to interpret the observed secondary task performance decrements. Norman and Bobrow (1975) introduced the important concepts of resource-limited processes, which are limited by the effort invested in a task and the priority placed upon task performance, and data-limited processes which are constrained by the quality of information rather than by increases in effort. In work situations, operators may compensate for any increase in task demands by increasing the amount of effort invested in the task. Therefore observed performance levels may remain constant but the operator experiences increased workload. Conversely, a reduction in the level of performance may result either because operators cannot maintain the level of effort expenditure required, or because they lower their criteria for adequate performance. Therefore task performance in resource-limited tasks may be limited by the effort put into the task (related to the priority an individual places upon performance), as well as the difficulty of the task. Secondary tasks have to be chosen carefully in order not to introduce structural interference with the primary task, however secondary-task performance measures can provide a more systematic technique for analysing interference in multiple-task situations than many of the other workload measures.

Observations of error and slips of action may provide useful data concerning the demands of different tasks, but a sound theoretical model of errors must be used to categorise different actions and activities. The recent work by Empson (1991) and Stager (1991) highlights promising approaches to the study of errors in air traffic control.

Workload In Applied Settings

An important point to be made is that there is variation in the way that people do tasks and therefore in the effects and consequences of what is termed workload. Prolonged active management of the resources required to meet task demands may lead ultimately to a deterioration in performance, but there may also be implications for short-term well-being and longer-term health. The experience of workload is thus unlikely to depend simply on task load, but rather on the interaction of task demands, how these demands are dealt with by an operator, and the level of performance achieved. Task demands are important but are mediated by effort and the priority placed on the particular tasks.

The level of control that operators are able to exert in complex systems is an important factor in the relationship between task demands, performance and well-being. Studies typically show an advantage for active control over passive control (Rasmussen & Rouse, 1981; Sheridan & Johansson, 1976). In other words, open-loop strategies, involving a greater degree of planning and broader understanding of the system as a whole, are seen to be more skilled and efficient

than the closed-loop mode. However, Umbers (1979) found that even experienced operators resort to the closed-loop mode when under high levels of workload or when unfamiliar situations or problems occur. This could be a cause for concern as remedial action may be applied once critical events have occurred rather than the more desirable situation in which impending catastrophic events are predicted at a time when something can be done to prevent them occurring.

Air traffic controllers have been found to vary their strategies according to task demand. Sperandio (1978) suggested that controllers take fewer variables into account as the traffic load increases. Similarly, in a study of process control, Bainbridge (1974) found that subjects under pressure used quicker, less accurate methods of finding data values. Although discretion to use open-loop control may normally be preferred by operators and lead to enhanced performance and safety, this discretion could be seen as a demand in itself imposed by the structure of the task or job.

Steptoe (1983) discussed a number of studies that show how the effort required to exert control over situations may result in more pronounced physiological responses. Increased physiological activation, indicated by higher levels of blood pressure and heart rate, was observed when subjects were engaged in effortful problem-solving or activity in a controllable situation. Thus, performance may be maintained at a desirable level (determined by personal and perceived organisational goals) but the effort required to deal with the demands is observed as costs in other psychophysiological systems (e.g., Frankenhaeuser, 1986; Mulder, 1986). Frankenhaeuser has demonstrated various changes in catecholamine and cortisol excretion with increased work demands. The patterning of these changes reflects active management of work and opportunity for control over work. Lundberg and Frankenhauser (1978) found that lack of control was associated with elevated physiological arousal during noise stress, but ratings of effort and the particular pattern of endocrine activity were found to be different depending on levels of performance. Increased ratings of effort, and higher levels of both adrenaline and noradrenaline were found for subjects who performed well in noise, but no changes were observed in the group whose performance was impaired. Generally it has been found that increased catecholamine excretion and lower levels of cortisol excretion and lower levels of anxiety are associated with active processing strategies linked with increased control and effort investment. Distress and both increased catecholamines and cortisol levels are associated with passive conditions or strategies (Frankenhauser, 1979; 1986). Using different physiological measures, Tattersall and Hockey (1990) identified different activities in a simulated flight engineer task which resulted in different cardiovascular costs and subjective ratings of effort and concern. Heart rate appeared to be associated with concern, particularly during activities such as landing and take off, but suppressed heart rate variability and increased subjective ratings of effort were associated with the requirement for problem-solving activity in different activities.

Thus, in order to understand the relationship between demanding situations and changes in performance, well-being and health, it is necessary to investigate changes in different domains. This involves the short-term and long-term assessment of individual state, both physiological and affective state (in terms of mood and well-being), as well as cognitive activity (as implicated in performance). An example of such a study was carried out to investigate the impact of naturally varying workload in air traffic control on a range of measures including performance, and physiological and affective state (Farmer, Belyavin, Tattersall, Berry & Hockey, 1991; Tattersall & Farmer, 1993).

Data were collected from 66 air traffic controllers during two whole working shifts at different times of the year. One shift was during the busy summer period from June to August

and the other during the relatively quieter periods, in terms of traffic load, in spring or autumn. Subjective ratings of mood (anxiety/tension, depression, and fatigue) were derived from a 20-item mood adjective checklist (Warr, 1989), which was completed at the beginning and the end of the shift and during each break period (usually every two hours during the shift). Subjective ratings of workload were collected at the same times throughout each day using the NASA Task Load Index (TLX) (Hart & Staveland, 1988). Saliva and urine samples were collected from which measures of cortisol and catecholamines were derived. ECG measures of heart rate and heart rate variability were collected continuously throughout the shift. The controllers also completed a set of cognitive tasks at the beginning and the end of each day. These tasks assessed visual vigilance and verbal reasoning and were used because of the difficulty in measuring overall performance in complex work systems such as air traffic control. Errors in performance may be infrequent and minor slips difficult to detect (Empson, 1991). Therefore tasks were chosen which were thought to simulate different aspects of air traffic control work, and to be sensitive to fatigue and levels of workload over the working day, for example in vigilance performance after prolonged periods of monitoring.

The analyses of these different measures show a number of interesting differences between and within high and low workload shifts indicating the negative consequences of dealing with sustained demands. Subjective ratings of workload were higher during the summer months. A simple measure of traffic count, however, did not appear to be an adequate predictor of subjective workload. Communications load was more clearly associated with the TLX dimensions of mental demand and effort, and traffic load with frustration.

The dimensions of mood were affected in different ways by increased workload. Anxiety showed a significantly greater increase during high workload days but pre-shift levels were not affected by workload. In contrast, levels of depression and fatigue were both higher at the start of the day under high workload conditions and were also elevated during the high workload shift. The sustained demands of the busy summer months appear to result in chronic after-effects of fatigue and depression, whereas anxiety was affected more transiently. However, repeated measurements are required at different times of day from off-duty days in addition to workdays in order to confirm this suggestion.

Salivary cortisol concentration was greater during high workload than low workload shifts, and greater before the shifts than later. It declined during the shift but not as clearly in the latter half of the shift under high workload. There were no reliable effects of workload in the analyses of urinary cortisol or urinary adrenaline but there was a significant interaction between workload and the two halves of the shift for urinary noradrenaline. Noradrenaline excretion decreased over the low workload shift but increased in the second half of the high workload shift. These findings perhaps reflect active coping with the quantity of demands during high workload shifts. The pattern of hormone excretion during high workload is consistent with the findings of Frankenhauser (1986) for example, in that cortisol and noradrenaline excretions are greater under conditions associated with lowered control and increased distress. This pattern may have long-term consequences for the health and well-being of controllers if sustained over long periods.

Heart rate did not differ reliably between high and low workload shifts but differences between the working periods, rest breaks and pre- and post-shift testing periods were found which may reflect differences in physical activity during the different periods. Increased heart rates were also found in the busier units (Heathrow Airport Control Centre and London Air Traffic Control Centre), but it is difficult to associate levels of workload or task variables with these differences as the factor of work location failed to interact with any of the other variables. Heart rate variability did not show a difference between high and low workload conditions,

although there was a significant effect of activity, indicating increased mental effort investment during the performance testing periods, compared to the working periods and rest breaks. At the level of analysis carried out to date the heart rate results are not very clear, but they confirm earlier results that heart rate variability is sensitive to effort expenditure in laboratory tasks (e.g. Mulder, 1980). However, this measure may show different patterns of effort expenditure for different individuals and the gross level of analysis of group data could be obscuring the more complex effects of workload.

Visual vigilance performance was affected by workload but verbal reasoning performance appeared to be insensitive to the effects of work between testing times, or between high and low workload shifts. For visual vigilance performance the signal detection measure of d' (sensitivity) was lower before the shift than after the shift, and there was a significant interaction between workload and time of testing, revealing improved sensitivity over low workload days but not during high workload days. No significant effects of workload or time of testing were found for the criterion measure (beta). Verbal reasoning performance was superior at the end of the shift compared to pre-shift performance. Speed of response did not differ significantly between high and low workload shifts, but was faster after the shift than before. A comparable analysis of errors showed no differences between high and low workload but errors were less frequent at end of the shift. These results support the notion that controllers become more actively engaged in their task during the shift, which is consistent with findings that the speed of unpaced work seems to increase later in the day (Broadbent, 1971). The performance of air traffic controllers tends to show an improvement over the day with the important exception of visual vigilance sensitivity, the ability to detect signals in noise. This measure showed an improvement on low workload days but not on high workload days, and suggests that heavy work demands in air traffic control may have a detrimental effect on monitoring performance.

Further analyses will investigate the role of individual differences in workload management by examining differences in the trade-offs between performance, affective state and physiological state as a function of style of coping and locus of control. It is hypothesised that more active copers will be more likely to maintain performance under high workload conditions but will show greater physiological and psychological effects of this activity. It would be desirable for further research to focus on the effects of sustained demands over longer periods of time than was possible in this study, as there are indications that there may be chronic effects of sustained exposure to high workload. Other research suggests that the morning shifts investigated in this study were not associated with the highest ratings of workload and fatigue (Farmer et al., 1990), and therefore the reported effects may actually be an underestimate of the potential impact of workload on well-being and performance.

It is argued that multi-level measurement techniques can provide a broad assessment of the impact of different work demands. Further studies, both controlled laboratory-based studies and field-based studies, are necessary to refine the techniques, but a model of stress and workload management based upon findings from such studies should allow the more accurate prediction of states or situations in which a breakdown in skills might occur. Such a breakdown is referred to by air traffic controllers as 'losing the picture', when they experience difficulties in attending to, and remembering accurately, relevant information about aircraft under their control. It is precisely this kind of situation that should be avoided in work in which safety is critically dependent upon performance.

Conclusions

The practical problems that might be encountered in carrying out workload evaluations in work settings have been outlined. Different approaches have been distinguished that may determine the type of research design used and provide assistance in the difficult choice between workload assessment techniques. One approach to workload assessment is to examine the short-term consequences of combining various tasks. Theoretical models of attention allocation (e.g. Wickens, 1984) will underpin specific studies of interference and the consequences of task demand and task conflict for performance. A further approach with a different temporal orientation may lead us to a better understanding of the relationships between work demands and strain through the analysis of individual differences in cognitive control processes. The application of these processes may depend on individual differences in long term styles and short term strategies, but may be used to prevent decrements in work performance under difficult conditions. However, control may attract costs as well as benefits in terms of changes in affective state and physiological activity. Thus, strain associated with work demands may only be measurable in the form of tradeoffs between performance and other domains of individual activity. The methodological implications are to identify patterns of adjustment to workload variations using repeated measures and longitudinal sampling of performance as well as subjective and physiological measures.

Possible enhancements to workplace design must take into account these human factors considerations of workload in order to avoid potential decrements in individual performance and associated organisational problems.

References

- Aasman, J., Mulder, G., & Mulder, L. J. M. (1987). Operator effort and the measurement of heart-rate variability. *Human Factors*, 29, 161-170.
- Bainbridge, L. (1974). Analysis of verbal protocols from a process control task. In E. Edwards & F. P. Lees (Eds.), *The Human Operator in Process Control*. London: Taylor and Francis.
- Beatty, J. (1982). Task-evoked pupillary responses, processing load, and the structure of processing resources. *Psychological Bulletin*, 91, 276-292.
- Broadbent, D. E. (1971). *Decision and Stress*. London: Academic Press.
- Damos, D. L. (1991). *Multiple-Task Performance*. London: Taylor and Francis.
- Empson, J. (1991). Cognitive failure in military air traffic control. In J. A. Wise, V. D. Hopkin & M. L. Smith (Eds.), *Automation and Systems Issues in Air Traffic Control*. Berlin: Springer-Verlag.
- Farmer, E. W., Belyavin, A. J., Berry, A., Tattersall, A. J., & Hockey, G. R. J. (1990). *Stress in Air Traffic Control I: Survey of NATS Controllers*. RAF Institute of Aviation Medicine Report No. 689.
- Farmer, E. W., Belyavin, A. J., Tattersall, A. J., Berry, A., & Hockey, G. R. J. (1991). *Stress in Air Traffic Control II: Effects of Increased Workload*. RAF Institute of Aviation Medicine Report No. 701.
- Frankenhaeuser, M. (1979). Psychoneuroendocrine approaches to the study of emotion as related to stress and coping. In H. E. Howe, & R. A. Dienstbier (Eds.), *Nebraska Symposium on Motivation*. University of Nebraska Press.
- Frankenhaeuser, M. (1986). A psychobiological framework for research on human stress and coping. In M. H. Appley & R. Trumbull (Eds.), *Dynamics of Stress*. New York: Plenum.

- Frankenhaeuser, M., & Johansson, G. (1986). Stress at work: psychobiological and psychosocial aspects. *International Review of Applied Psychology*, 35, 287-299.
- Gopher, D., & Donchin, E. (1986). Workload – An examination of the concept. In K. R. Boff, L. Kaufman & J. P. Thomas (Eds.), *Handbook of Perception and Human Performance Volume II: Cognitive Processes and Performance*. New York: John Wiley & Sons.
- Hancock, P. A., & Meshkati, N. (1988). *Human Mental Workload*. Amsterdam: Elsevier.
- Hart, S. G., & Staveland, L. E. (1988). Development of a NASA TLX (Task Load Index): Results of empirical and theoretical research. In P. Hancock & N. Meshkati (Eds.), *Human Mental Workload*. Amsterdam: Elsevier.
- Hill, S. G., Iavecchia, H. P., Byers, J. C., Bittner, A. C., Zaklad, A. L., & Christ, R. E. (1992). Comparison of four subjective workload rating scales. *Human Factors*, 34, 429-439.
- Hulbert, T. (1989). A comparison of the 'NASA-TLX' and 'ISA' subjective workload rating techniques. Civil Aviation Authority Air Traffic Control Evaluation Unit, Bournemouth, UK. Internal Report.
- Jordan, C. S. (1992). Experimental study of the effect of an instantaneous self assessment workload recorder on task performance. *Defence Research Agency Technical Memorandum DRA TM (CADS) 92011*. Portsmouth, Hants: DRA.
- Kramer, A. F. (1991). Physiological metrics of mental workload: A review of recent progress. In D. L. Damos (Ed.), *Multiple-Task Performance*. London: Taylor & Francis.
- Lundberg, U., & Frankenhauser, M. (1978). Psychophysiological reactions to noise as modified by personal control over noise intensity. *Biological Psychology*, 6, 55-59.
- Mackay, C., Cox, T., Burrows, G., & Lazzarini, T. (1978). An inventory for the measurement of self-reported stress and arousal. *British Journal of Clinical Psychology*, 17, 283-284.
- Moray, N. (Ed.) (1982). *Mental Workload: Its Theory and Measurement*. New York: Plenum.
- Moray, N. (1988). Mental workload since 1979. *International Reviews of Ergonomics*, 2, 123-150.
- Mulder, G. (1980). *The Heart of Mental Effort*. University of Groningen, The Netherlands.
- Mulder, G. (1986). The concept and measurement of mental effort. In G. R. J. Hockey, A. W. K. Gaillard, & M. H. G. Coles (Eds.), *Energetics and Human Information Processing*. Dordrecht: Nijhoff.
- Norman, D. A., & Bobrow, D. G. (1975). On data-limited and resource-limited processes. *Cognitive Psychology*, 7, 44-64.
- Nygren, T. E. (1991). Psychometric properties of subjective workload measurement techniques: Implications for their use in the assessment of perceived mental workload. *Human Factors*, 33, 17-33.
- O'Donnell, R. D., & Eggemeier, F. T. (1986). Workload assessment methodology. In K. R. Boff, L. Kaufman & J. P. Thomas (Eds.), *Handbook of Perception and Human Performance Volume II: Cognitive Processes and Performance*. New York: John Wiley & Sons.
- Ogden, G., Levine, J., & Eisner, E. (1979). Measurement of workload by secondary tasks. *Human Factors*, 21, 529-548.
- Rasmussen, J., & Rouse, W. B. (Eds.). (1981). *Human Detection and Diagnosis of System Failures*. New York: Plenum.
- Ravden, S., & Johnson, G. (1989). *Evaluating Usability of Human-Computer Interfaces: A Practical Method*. Chichester: Ellis Horwood.
- Reid, G. B., & Nygren, T. E. (1988). The subjective workload assessment technique: A scaling procedure for measuring mental workload. In P. A. Hancock & N. Meshkati (Eds.), *Human Mental Workload*. Amsterdam: North-Holland.
- Rehmann, J. T., Stein, E. S., & Rosenberg, B. L. (1983). Subjective pilot workload assessment. *Human Factors*, 25, 297-307.

- Rohmert, W. (1987). Physiological and psychological work load measurement and analysis. In G. Salvendy (Ed.), *Handbook of Human Factors*. New York: John Wiley & Sons.
- Sheridan, T. B., & Johanssen, G. (Eds.). (1976). *Monitoring and Supervisory Control*. New York: Plenum.
- Sperandio, J. (1978). The regulation of working methods as a function of workload among air traffic controllers. *Ergonomics*, 21, 195-202.
- Stager, P. (1991). Error models of operating irregularities: Implications for automation. In J. A. Wise, V. D. Hopkin & M. L. Smith (Eds.), *Automation and Systems Issues in Air Traffic Control*. Berlin: Springer-Verlag.
- Stephoe, A. (1983). Stress, helplessness and control: The implications of laboratory studies. *Journal of Psychosomatic Research*, 27, 361-367.
- Tattersall, A. J., & Farmer, E. W. (1993). The regulation of work demands and strain. In G. P. Keita, & S. L. Sauter (Eds.), *Job Stress 2000: Emerging Issues*. Washington, D.C.: American Psychological Association.
- Tattersall, A. J., & Foord, P. S. (1993). An experimental evaluation of instantaneous self assessment as a measure of workload. Submitted to *Ergonomics*.
- Tattersall, A. J., & Hockey, G. R. J. (1990). The assessment of workload in a complex monitoring and fault diagnosis task. In D. Brogan (Ed.), *Visual Search*. London: Taylor & Francis.
- Umbers, I. G. (1979). Models of the process operator. *International Journal of Man-Machine Studies*, 11, 263-284.
- Vicente, K. J., Thornton, D. C., & Moray, N. (1987). Spectral analysis of sinus arrhythmia: A measure of mental effort. *Human Factors*, 29, 171-182.
- Vidulich, M. A., & Tsang, P. S. (1987). Absolute magnitude estimation and relative judgement approaches to subjective workload assessment. In *Proceedings of the Thirty First Annual Meeting of the Human Factors Society*. Santa Monica, CA: Human Factors Society.
- Warr, P. B. (1989). The measurement of well-being and other aspects of mental health. *Journal of Occupational Psychology*, 63, 193-210.
- Wickens, C. D. (1984). Processing resources in attention. In R. Parasuraman & D. R. Davies (Eds.), *Varieties of Attention*. New York: Academic Press.
- Wickens, C. D. (1991). Processing resources and attention. In D. L. Damos (Ed.), *Multiple-Task Performance*. London: Taylor & Francis.
- Wierwille, W. W., & Casali, J. G. (1983). A validated rating scale for global mental workload measurement applications. In *Proceedings of the 27th Annual Meeting of the Human Factors Society*. Santa Monica, CA: Human Factors Society.
- Yeh, Y. Y., & Wickens, C. D. (1988). Dissociation of performance and subjective measures of workload. *Human Factors*, 30, 111-120.

