AD-A092 840 MASSACHUSETTS INST OF TECH CAMBRIDGE MAN-MACHINE SYS--ETC F/G 5/8
HUMAN INFORMATION PROCESSING AND SUPERVISORY CONTROL.(U)
MAY 80 N MORAY N00014-77-C-0256
UNCLASSIFIED NL

1 OF 1
AD
A092840

END
DATE
FILMED
1-81
DTIC

Man-Machine Systems Laboratory

Department of Mechanical Engineering
Massachusetts Institute of Technology
Cambridge, Massachusetts 02139

UNCLASSIFIED
SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER	2. GOVT ACCESSION NO. AD-A092 840	3. RECIPIENT'S CATALOG NUMBER
4. ~~TITLE (and Subtitle)~~ HUMAN INFORMATION PROCESSING AND SUPERVISORY CONTROL		5. ~~TYPE OF REPORT & PERIOD COVERED~~ 1 Jul[redacted]79 - 30 Jun[redacted]80
		6. ~~PERFORMING ORG. REPORT NUMBER~~
7. ~~AUTHOR(s)~~ Neville Moray		8. CONTRACT OR GRANT NUMBER(s) N00014-77-C-0256 ✓
9. PERFORMING ORGANIZATION NAME AND ADDRESS Massachusetts Institute of Technology Cambridge, Massachusetts 02139		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS NR 196-152
11. CONTROLLING OFFICE NAME AND ADDRESS Engineering Psychology Programs Office of Naval Research (Code 455) Arlington, VA 22217		12. ~~REPORT DATE~~ May [redacted]80
		13. ~~NUMBER OF PAGES~~
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)		15. SECURITY CLASS. (of this report) UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE
16. DISTRIBUTION STATEMENT (of this Report) Approved for public release; distribution unlimited		
17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)		
18. SUPPLEMENTARY NOTES NONE		

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

information	man-machine systems
human performance	monitoring
supervisory control	memory
	human operator

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

Although there are several well developed models of man as a manual controller, there is at present no general theory of supervisory control. However, with increasing use of microprocessors and automatic machinery the role of human operator is increasingly one of supervision rather than control, at least as long as the controller processes remain within the acceptable tolerances of normal performance.

The design of complex automatic systems which must be supervised by a

(OVER)

DD FORM 1473 1 JAN 73 EDITION OF 1 NOV 65 IS OBSOLETE
S/N 0102-LF-014-6601

UNCLASSIFIED
SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

human operator should benefit from considering the particular properties of man as an information processor which are likely to render supervisory control a difficult task. This is especially true due to the fact that while for long periods he may be required to do little other than observe the system, he must nonetheless extract enough information about the state variables to enable him to make adequate decisions, and exercise adequate control, if emergency situations develop.

This paper is an introduction to the empirical literature most relevant to matching the properties of the human operator to those of the machine when designing systems to be run under supervisory control. The literature citations are not intended to be exhaustive, but to be sufficient to establish the most important points which will be made.

CONTENTS

TAXONOMIES OF ABILITIES, DEMANDS, TASKS, & FUNCTIONS

Although the vast majority of experimental studies of man as a processor of information have been carried out in psychological laboratories, the situations with which I shall be concerned in this paper call for a different standpoint. It is probably that most of the most sophisticated experimental studies and the theories developed from them by psychologists may not be directly applicable to the design of man-machine systems of any complexity. Rather, it will be necessary to select certain studies which, despite the differences between laboratory studies and the industrial situation, can be used to outline the main problems which are likely to be encountered in the design of man-machine systems in the areas of process control, power stations, aircraft cockpits, remotely controlled underwater vehicles, etc.

This chapter outlines the main categories which will be used throughout the paper. But we may conveniently begin by looking at some recent remarks by Rasmussen (1979) about the difference between laboratory and industrial behavior of the human operator. He remarks,

> "Laboratory tasks tend to have a well defined goal or target. Payoff matrices are artificial and have low values. The subject is controlled by the task. Task instructions are specific. Task requirements are stable. Subjects are relatively untrained. By contrast in "real" tasks only a (sometimes vague) overall performance criterion is given and the detailed goal structure must be inferred by the operator. Task instructions are inferred by the human operator from rather general commands about how to perform the task. The task may vary as the demands of the system vary in real time. Operating conditions and the system itself are liable to change. Costs and benefits may have enormous values. There is a hierarchy of performance goals. The operator is usually highly trained, and largely controls the tasks, being allowed to use what strategies he will.. Risk is incurred in ways which can never be simulated in the laboratory."

The world of the psychological laboratory is very different from the world implied by Rasmussen's description. The typical paradigm of psychological research involves not more than two or three variables, and often only one. An operator who invents ways of carrying out the task different from those specified by the experimenter will be rejected, even if they are more efficient. The effects of practice are reduced to a minimum. Variables are kept for the most part statistically independent of one another. The duration of the experiment is generally short, not more than a few hours at most. Trials are statistically independent. The behavior of the operator has no effect on what happens next. There are, obviously, exceptions to this description of a classical investigation in experimental psychology, but it is typical of the vast majority of studies.

Properties of the World

By contrast, consider the properties of the world in which the operator of a complex man-machine system typically works.

1. The world is complex. It is characteristic of "real" tasks that there are many sources of information whose values must be simultaneously known if the current state of a system is to be well defined.

2. The world is dynamic. As Heraclitus observed you cannot even step into the same river once, let alone twice. The world has a history which unfolds in the direction of time's arrow, and the most important information about the world is information about change. As Kelley (1968) said, the reason that an operator requires status information about a system is in order to predict what is about to happen, not to know what is now happening. The dynamics may be external, in the sense of the forcing functions of a system or injected disturbances, or internal, in the sense of changing transfer functions of the controlled elements. Adaptive behavior, whether in the natural world or the artificial world of the factory or aircraft cockpit must be behavior appropriate to a dynamic process, not to the occurrance of a particular stimulus.

3. The role is bandlimited. Most of the processes with which we will be concerned are relatively slow processes. There may be occasional emergencies which call for responses in a fraction of a second, but they are rare. This may be simply because there is an intuitive realization that humans cannot control processes with high bandwidth, and therefore design has evolved in such a way that either processes run slowly or fast systems are always automatically controlled. There are, however, exceptions - a high performance aircraft during the final phases of landing, driving an automobile at speed in traffic, or the events during the first few seconds of a failure in a nuclear power plant (Spectrum, November 1979). By and large organisms live at the low frequency end of natural events.

4. The world is noisy. Phenomena such as fog, glare, lighting, acoustic noise, etc., means that the human operator must interpret incoming sensory information by means of some kind of statistical decision. All sensory information is more or less ambiguous. Even when good design makes the sensory information as clear as possible at the display, the nervous system is noisy and renders perceptual decisions statistical in character. The aim of the design of instruments should

be to render the signal to noise ratio as high as possible in all cases, taking into account both environmental variables and what is known about the psychology and physiology of the human operator.

5. The world is somewhat Markovian. In general it is possible to make reasonable estimates of the immediate future states of the world from present observations. These predictions are only stochastic, but this is enough in principle to mean that the human operator can, in principle, develop tactics for the (near) future on the basis of present observations. Whether he does so efficiently is, of course, another matter.

6. The world is rather closely coupled. In general information from adjacent regions in time and space are related, and significantly correlated. The world is not the set of independent random generators beloved of experimental design in psychology laboratories, and knowledge about one part of it usually contains useful information about neighboring regions. This is especially true of man-made systems. since they are designed to embody causality, and their complexity resides precisely in the large number of closely coupled elements. This does not mean, however, that conrolling one part will always help in controlling another part. Since closely coupled regions may have different control laws and require behavior which is non-homogeneous and which may cause mutual interference. It should be noted that with the development of electronic and other communication channels the concept of "near" takes on a new meaning, since parts of a system may be functionally near but physically extremely distant.

7. The world is meaningful to humans. Very few events in the world are treated by human operators as merely events, or "stimuli". Rather they are a source of hypotheses, consciously or unconsciously evaluated, about the structure of causality. Man approaches the task of deciding on action with reasons, not merely responses (Taylor, 1976). This can result both in enormous economics in information processing when the hypotheses are correct, and highly maladaptive behavior when they are wrong (Weiner, 1977).

8. The world has values for humans. Events in it are more or less important, even though they may have equal amounts of information in a statistical sense. Consequently the adaptive response by the operator to states of the world cannot be accounted for without understanding his value hierarchy, his assessment of costs and payoffs, which may be either objectively imposed by the system specifications, or subjectively estimated by the operator, or both.

The above description is intended to be true either of the natural world or of the world as encountered in complex man-machine systems. Within the latter, however, it is possible to describe a brief taxonomy which covers most of the tasks which are likely ot be encountered in the near future in supervisory systems. These are:

1. The control of continuous variables.
2. Monitoring continuous variables.
3. The control of digital variables.
4. Monitoring digital variables.
5. Tasks involving cognitive decision making (such as Air Traffic Control).
6. Problem solving tasks (such as circuit fault diagnosis).
7. Failure detection, fault detection, and error detection.

In supervisory control (7) is perhaps the central tasks, along with monitoring, since under normal conditions the system itself will perform the control operations. Failure detection means the detection of a permanent change in system behavior such that the output variables will stay outside their desired operational range for long enough to require intervention beyond the capabilities of the system by itself. Continuous variables are those displayed by an analogue display, however generated, while digital variables are those which are displayed as numbers or discrete symbols. Analogue control uses joysticks, wheels, etc., while digital control uses a keyboard or other discrete input device which exercises control commands using symbolic information, whether numerical or logical. Mixtures of the different types abound. In Air Traffic Control, for example, there is an analogue display (radar), cognitive decision making, and control may be exercised either through keyboard entries or verbal communications.

Properties of the Human Operator

Finally, for the purpose of organizing the discussion and without intending it to be a formal and exhaustive description of the human operator, we may list the main operations which the human must perform when acting as a supervisor or controller.

1. Intake of information which is rendered difficult by the fact that the visual system (which is the usual channel through which accurate information is received) is highly directional and the sources of information may be spread over a large area. The fact that the sources are noisy and the peripheral nervous system is noisy further complicate data acquisition.

2. Interpretation of information which includes the process of identifying the signals (consciously or unconsciously) and which is difficult because of the noisy nature of the data. The data may also be inadequate because less than an optimal amount of data was received, either due to the dynamics of the source or display, or because the operator did not sample the source for an appropriate duration. In general the interpretation of data involves a statistical decision about the nature of the information received from (1).

3. Decision making which is rendered difficult by factors already mentioned. The decisions include those which lead to perceptual classification, decisions to do with tactics for the further sampling of the sources of information, and decisions to do with the initiation of overt action and interaction with the system being supervised.

4. Incorporation of information into permanent storage. This rather clumsy phrase is used to emphasize that there is more than one way of retaining the effect of past experience. On the one hand information may be stored in memory in the form of verbal or visual tactile or kinaesthetic images, which can be accessed voluntarily and used in consciousness. On the other hand, skilled actions may be learned, and be neither describable nor imaginable, but exist as performance abilities. In addition, there are properties of learning such as generalization, by which information which has been identified is used as if it were other information which has been closely associated with it in the past; and also the profoundly important concept of an "internal model" which is a more-or-less accurate representation of a subset of the features of the environment which can be used by humans as a substitute for input when predicting the future course of events.

5. Extrapolation and prediction by which present (or recent) information is used to predict future states of the environment on the basis of past experience during training and operations.

6. Generation of information and action which includes both overt behavior which results in a change in the environment (including the controlled elements of the physical plant), or in the relation of the human operator to the controlled elements; and also the sending of verbal messages. In addition there are responses which remain covert and do not emerge directly as behavior, but which modify the operator's future behavior, such as the choosing of tactics (such as a different equalization law in the face of variations in plant characteristics or environmental status), changes in sampling behavior which may not become apparent for some time, etc. The output mechanisms, by which include both the central mechanisms which choose what actions to perform and also the muscle control mechanisms, appear to be rather more error prone than has been thought. Most estimates have used measures of performance errors (that is of the output of the man-machine system), not of behavior errors (that is of the output of the human operator). There is growing evidence (Senders, personal communication; Norman, personal communication; Ruffle-Smith, 1979) that behavioral errors are quite frequent in the control of complex systems. However, the human operator appears also very adept at noticing the errors which he has made, and in correcting them before the performance errors become large enough to be noticeable.

Confronted by a world with the properties listed above, and provided with the information processing abilities which limit him in various ways, how should a human operator behave in order most effectively to carry out the task of supervisory control? In any particular case the "world" will be the variables which are displayed by the system which he is monitoring and controlling, and his abilities may be hindered or enhanced by suitable design of the man-machine interface. We now turn to the relevant evidence from experimental psychology.

SOURCES OF DIFFICULTY IN SUPERVISORY CONTROL

In considering what aspects of man-machine system design may cause difficulty for the human operator engaged in supervisory control, it is convenient to divide the problem into two phases. Normal supervisory control is essentially a search task. The set points of the many variables are specified, and tolerable errors around these set points are known to the operator. There may be a very large number of variables involved, but essentially the task of the supervisor is to scan the sources of information in a suitable way to ensure that their values are within the acceptable error tolerances. With conventional instrumentation this phase of control requires very little action on the part of the supervisor. The system is assumed to be an automatic one (otherwise we would not call it supervisory control, but manual control) and until an abnormal state occurs there is no call for any control action.

This may not be true in future with new kinds of displays being developed. For reasons which will be discussed below it appears to be widely felt that conventional instrumentation should be replaced by "integrated displays", or even by displays with computer-generated graphics on which the operator may call up any subset of the total state variable set which he wishes. He now has a rather different, and considerably more active role to take in monitoring the system. His actions will not directly affect the system; that is, he is still not a controller, but in order to examine a state variable he must call it up onto his display. It is true that no longer need he walk about the control room, but action is required, and inappropriate action will now render his search less than optimally effective. Hence from matters of tactics and strategy which will be discussed below, straightforward questions of ergonomics enter at this point into the design of graphic displays, keyboards, etc. It is fair to say that less than adequate consideration of human factors has been given to this question. (For example, a widely used keyboard has a 'no scroll' key in such a position that an operator resting his hands for a moment with the wrists on the table can easily hit the 'no scroll' key inadvertently, thus making the screen inoperative without there being any indication that this has happened).

On the other hand, when the supervisor detects an abnormal situation he must once again become an active controller. The exact mode of control will depend upon the system. The failure of an automatic pilot in an aircraft may necessitate the pilot changing from a flight manager to a classical manual controller using analogue controls (joystick, throttle, rudder, etc.). Failure of control in a process control plant may call for the activation of valves, the use of switches, and in modern computer controlled systems for data entry and command entry on the keyboard or other form of data entry device (light pen, touch tablet, etc.).

We may therefore divide the discussion of the sources of difficulty in supervisory control into two sections. As we shall see later, some of the sources of subjective task difficulty (high order of manual control, plant instability in manual control, nonhomogeneity of control laws in multidimensional control, etc.) are not applicable during normal plant operation, but may become relevant during failure. But other sources of difficulty, such as signal-to-noise ratio, memory load, and above all time stress apply as much if not more in supervisory control as in manual control, and especially with new instrumentation. It should be noted that there is very little direct research on supervisory control. Most of what follows is extrapolation from more conventional studies, and may be taken partly as suggestions for research.

A Scenario

An operator sits at a console watching a large number of sources of information. The console contains an array of analogue and digital displays, from which he may ascertain the value of the state variables of the system. Some of these will have variable set points, and he must make sure that the actual value of the variable does not depart 'significantly' from the desired value. The normal operation of some sub-systems is indicated by the position of switches, or by lights, which may be color coded. He may interact with the system by means of switches, thumbwheels, etc. In addition, he has one or more visual display units (VDU's) which display all or part of the system as a flow chart of mimic diagram, and he can call for the value of variables to be displayed on the VDU's. This he does by means of a keyboard for data entry, or by a light pen or touch tablet. The total display subtends about 120° visual angle in the horizontal dimension and about 60° vertically.

What will be his chief sources of difficulty in carrying out his task as a supervisor, in the light of the known properties of the human operator as a processor of information, and what optimal features of design, or strategies, will aid him?

We begin with some general comments. Certain of the properties of the world which were outlined earlier will help, and certain hinder the operator. (The world is defined as the system he is controlling and all the variables which affect him and it.) The fact that the system is complex is a source of difficulty, since the complete diagnosis of the system state will require information from many sources of information. In general, a particular output will have several possible causes. The complexity of the systems has led to the increasing popularity of a state space representation, but there have been almost no psychological experiments carried out using such framework, because of the desire for experimental simplicity.

The fact that several sources of information must be sampled gives rise to two sources or workload. The first is due to man's being functionally a single channel system in such situations. Because of this it will take a long time to sample several channels, and during this period critical events may occur on channels which are not being examined. This will be increasingly true as the size of the display increases, and in large power stations the number of displays and controls may run into the thousands. Warning lights or sounds may help to capture the attention, but these are not themselves always reliable, and if many are activated simultaneously there will be a diagnostic problem anyway. Large panels produce a load on the operator simply through their architecture. It is usually assumed that integated displays, or VDU's, will ease the operator's task by reducing the amount of bodily movement or eye movements required to scan the variables. The second source of load will be due to the necessity of retaining in memory the values of the variables which are examined. It is well known that the number of items which can be retained under normal conditions in immediate memory ("short term memory", "primary memory", "working memory") is about seven. This is not a large number, and performing any mental operations on the contents of memory tends to reduce the span. Running memory, in which new items are added and old items are dropped continuously is far smaller, generally not more than three items. Hence the need to combine information can be expected to be a severe source of difficulty in supervisory control.

The fact that the information reaching the decision making levels of the nervous system is noisy will interact with the rate at which the operator can accumulate information about the state vector, since in general noisy signals must be sampled for longer or more often than clear signals. This in turn will limit the rate at which multiple sources can be sampled.

The fact that the world is dynamic means that no source of information can be regarded as having a fixed value. Whatever strategy the observer adopts must be one which guarantees that all sources are sampled at a rate related ot their bandwidth, but modulated, as we have just seen, by the time taken to acquire accurate information when a sample is taken. The aperiodicity of the sampling pattern imposes a memory load on the operator, who must keep track of the sampling schedule. This may be a conscious or an unconscious process.

On the other hand the fact that the world is bandlimited and Markovian suggests that the operator should be able to take advantages of spatial and sequential correlation among the variables to economize on his information processing. Some prediction, both of simultaneous and of successive values, will be possible. The close coupling of parts of the system may also assist in this, although if active control is being exercised the coupling of nonhomogeneous control laws will increase task difficulty.

Meaning and value play an important part in developing strategies and tactics of behavior. The fact that the operator can understand the underlying causal structure of the system allows him to make hypotheses which are much more economical than the mere correlational relations among the values of variables. He may know, for example, that a fault in a particular part of the system necessarily must cause faults in certain other parts, without the necessity for examining the values of the displayed variables of the latter. Payoffs, costs and values may suggest biassing his sampling behavior in favor of more important variables. This in its turn may have the effect of changing a situation where several events are equally probable to one which may be treated as if some of them are less probable, and the resulting imbalance of "functional" probability, or bias, will reduce the rate of processing information. Often the understanding which the operator has of a system will enable him effectively to call up new "subroutines" of behavior and tackle the task in a different way, not merely trim parameter values in his existing control laws. It is interesting that one may find operators who claim to prefer older, less automatic systems, with "poorer" displays, because they claim that with the older systems they felt that they were controlling the process itself, whereas with the new displays and fully automatic systems they feel that they are merely controlling the positions of the indicators on the instruments, not the physical process which those indicators mirror. Keeping in direct touch with the physical process may be of great importance to an understanding of the process, and perhaps also of morale and motivation.

For practical purposes the human operator may be considered to be a single channel information processor with limited capacity. There are experiments which suggest that under certain laboratory conditions information intake may be in parallel. A number of workers have argued that all incoming sensory information is fully analysed up to and including pattern analysis, but not at the level of consciousness, and that prior to consiciousness the analysis proceed in parallel (Posner and Klein, 1973; for example). However these laboratory experiments require very special conditions, and certainly do not include cases where substantial eye movements are required to ensure accurate visual information processing. Because of the structure of the eye, and the need for foveal fixation for accurate pattern perception and the reception of color, man must function as a single channel system when scanning large arrays of instruments. The only useful information which can be obtained from the periphery of vision is rate of change information.

Similarly, while it is often said that when perceptual motor skills become automatized processing is automatic and does not impose a load, it is remarkable difficult to find experimental confirmation of that suggestion. Furthermore, while multiple axis joysticks may allow the operator to control more than one dimension simultaneously, there are many data entry devices - light pens, keyboards, etc. - which constrain his actions to single channel.

To a good practical working approximation man can, then, be regarded as both single channel and limited in capacity. In this context "single channel" means "processes only one input or output at a time regardless of what its information content is", and "limited capacity" means that there is a maximum rate at which signals can be processed even if only one source is sampled. A single channel system will not pass several signals at the same time even though their total information content is less than the channel capacity. (While this interpretation is not strictly in line with Shannon's original formulation of information transmission theory, it is the common usage in psychological literature.) Conscious decision making mechanisms are certainly single channel, in that we are only able to think of one thing at a time.

In so far as man is single channel and has a limited memory, the contention of Senders (personal communication) that all load is due to time stress has some basis. It is rather obvious that an Air Traffic Controller trying to handle ten aircraft at once is overloaded, and that the time taken to service each aircraft is a crucial variable. But it is not so obvious why, say, single axis compensatory tracking at a bandwidth of 1.0 Hz is time stressing, where the forcing function is a zero mean Gaussian noise. Indeed there _is_ no time stress in such a task because no required degree of precision has been stated. But as soon as a permissible error, say ± _e_ rms, is specified the situation changeds. The operator _must_ now make corrections more frequently as _e_ is set to smaller and smaller values, for if he does not the function will depart from zero by more than the permissible error. Even at low bandwidths a sufficiently small value of _e_ will call for frequent inputs from the operator. Thus precision imposes a time stress. Moreover, the transfer function of the controlled element may be such as to introduce a phase lag or delay, so that the effect of the operator's response does not become apparent for some time, and during that delay the error will increase again. If the controlled element is unstable, then frequent corrections are mandatory. Hence even a continuous low frequency task can have properties which may apear as time stress to a single channel system with poor memory.

On the other hand, if we consider a task such as fault tracing in an electronic circuit or debugging a computer program, there may be no real time stress other than that imposed by the operator on himself. (This would not, of course be true if the fault tracing or debugging was being done in real time online to an automatic system which was running in the meantime; there the time stress would again be obvious.) Most people have some limit on the time for which they will try to solve a problem. If given an insoluble problem, one may guess that one of two things will happen. Either the person will reach his subjective time limit and simply admit that he cannot solve it; or he would invent an arbitrary solution which would allow him to stop. (For example, given a short series of random numbers and asked for the next one, he might, in an effort to find a pattern, eventually say "the series repeats again".) It is well known that neophyte computer programmers frequently decide that there must be a hardware fault in the system after they have tried for what they subjectively feel is a "reasonable time" to find the error in the program. The length of time considered "reasonable" in any particular situation is a matter of acquired social norms and will be affected by payoffs, etc., and if payment is involved in terms of "piece rates", than a very obvious time stress will arise.

There is another sense in which time stress arises in continuous tasks as tracking, or monitoring continuous functions. Sheridan (1970) has discussed the problem of the optimal sampling interval for a supervisor. If we assume that the operator can extrapolate the future course of the function from an observation, and also that there is a cost associated with each observation, then an optimal sampling interval can be specified. The calculation of this interval, and the estimation of its expiry, then become a secondary task which imposes stress, for the estimation must be completed before the interval is over.

The fact that information is rapidly lost from memory, either due to forgetting or due to interference from new information, adds a further time stress, for if information from different sources must be combined, then clearly it is desirable to do so as rapidly as possible. If information is acquired froma noisy source cumulatively, a problem arises due to the fact that the earlier samples will be fading while the later samples are being gathered. There should be an optimal interval at which to make a decision, calculable from the convolution of the data acquisition curve and the forgetting curve, and either earlier or later decisions should be less accurate. AT least two papers (Megan dn Richardson, 1979; Wickens and Kessel, 979) suggest that such an optimal duration exists.

If we now add the constraint that there is more than a single source which must be examined, the presence of time stress is immediately apparent unless the signals are perfectly correlated. A single channel observer must finish processing information from one source as rapidly as possible if he is to avoid the risk of missing significant events on the other source.

The degree of time stress is largely determined by the speed accuracy tradeoff, which in its turn is set by the operator's understanding of the task demands and the demands of the instructor who has set him his overall goal. Tasks in the real world must be finished in a reasonable length of time and to a reasonable degree of accuracy, and the emphasis of time suggested by Senders is perhaps better seen as an emphasis on satisfying the moment to moment demands of the speed accuracy tradeoff. Workload in supervisory control systems arises because decisions have to be made at a satisfactory rate and at a satisfactory degree of accuracy.

Intake and Interpretation of Information

Sampling Strategies

If the human operator is well approximated as a single channel observer, how is his sampling schedule controlled? In the absence of any other information, the best solution is to behave in accordance with the Sampling Theorem of Shannon and Gabor. A source which is bandlimited at W Hz should be sampled 2W times a second. It has been known since the work of Senders, Grignetti, Elkind and Small wood (1966) that under some conditions humans do behave in such a way. They predicted the distribution of attention over four instruments with different bandwidths from the Sampling Theorem when attention was defined in terms of eye movements. They also found that the duration of fixations was related to the information content of the signals, but did not report any interaction between the duration and frequency of fixations. At the bandwidths used there would have been ample time to read the signals accurately before a new fixation was required. Senders also reported that the eye movements of pilots were in good accordance with the theorem, and less well with the introspections of the pilots. It is interesting to note that his findings can be readily repeated in the laboratory, and that when the operators as well practiced the same results are obtained. However, early in practice quite different and very varied patterns of eye movements may be observed. For example, a beginner at the task may simply fixate each instruments in turn, going clockwise round the display. Moreover Hamilton (personal communication) in a task which is quite closely related to that used by Senders did not find the same result, and has suggested that the operator must be working near his limit for them to occur. If there is some spare "effort" available (to use Kahneman's (1973) popular but unilluminating term) the operator will develop conscious strategies, which will not match the optimal one defined by the Sampling Theorem. Moray, Synnock & Richards (1973) also found a situation where sampling was not in accordance with the theorem. In their case the rate of presentation was under the control of the operator, and a fixed interval sampling strategy developed. This could be interpreted to mean that the rate was varied to produce a bandwidth which allowed fixed interval sampling. At least in some tasks (Monk, 1977) fixed intervals are used more efficiently the human operator than irregular intervals. If a periodic sampling is optimal, conscious choice of intervals will be inefficient. Only long practice which results in the sampling becoming automatic will lead to efficient behavior in accordance with the sampling theorem.

The fact that all the operators investigated by Senders et al. ended by showing very similar behavior is significant. It implies that they had acquired an internal statistical model of the task structure, and that the model was guiding their behavior (see Figure 1, after Senders et al.) Their observers monitored four voltmeters which were driven by white noise of differing bandwidths. If a voltage greater than a specified criterion was observed the observers were instructed to press a button - the "response" in Fig. 1. In such an experiment widely differing

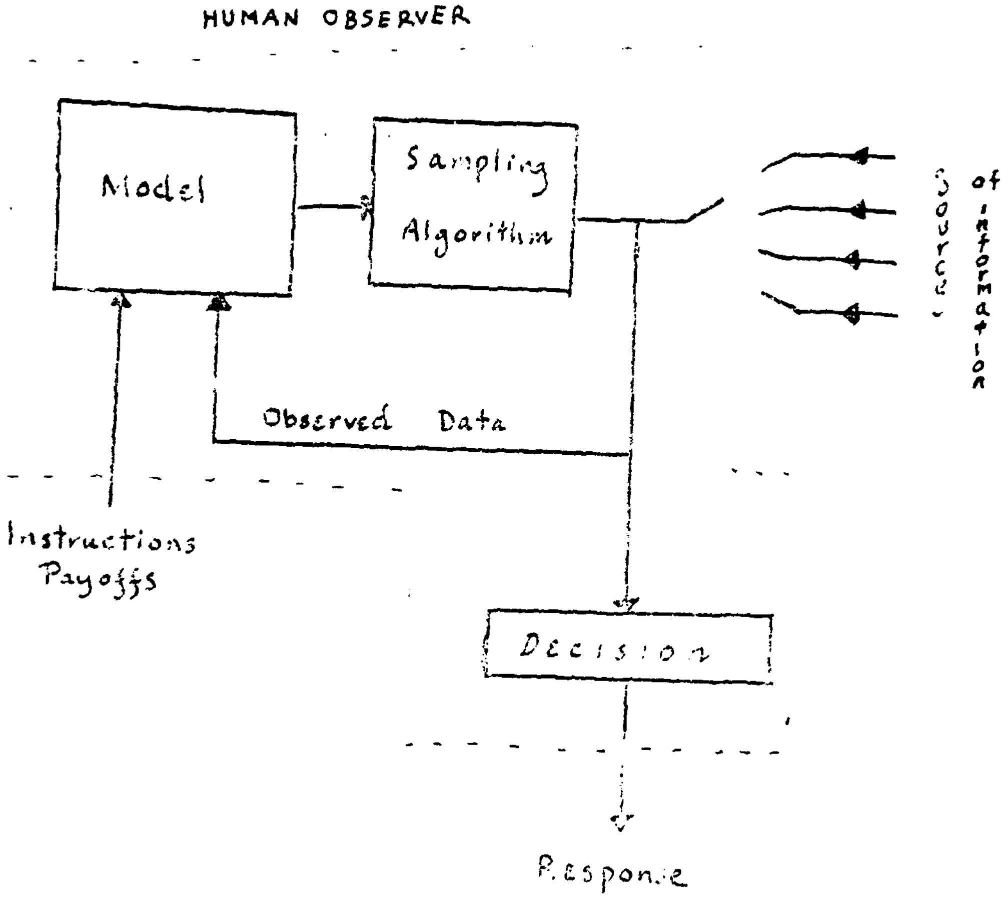


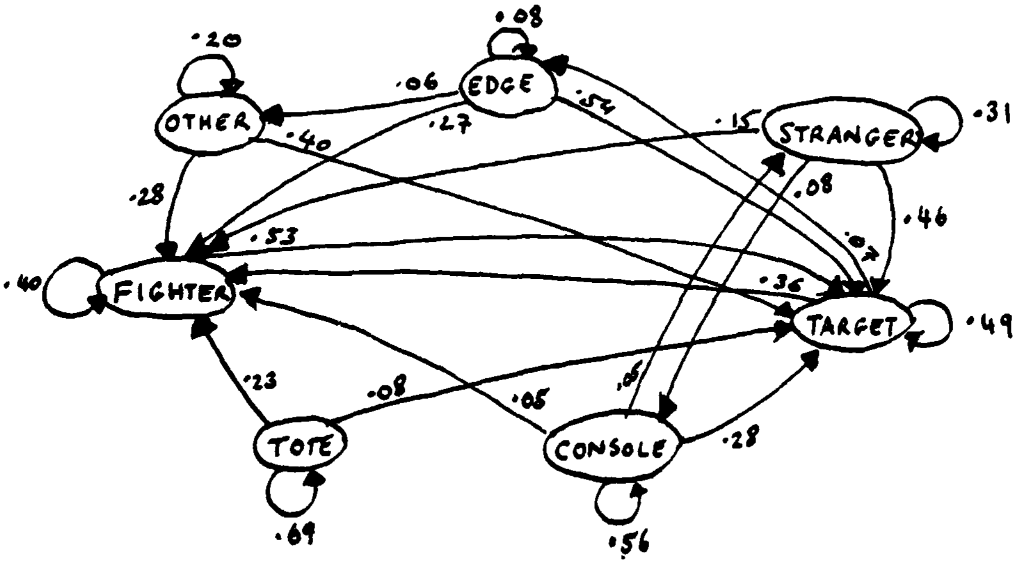
FIGURE 1: SENDERS' MODEL OF THE HUMAN OBSERVER

TABLE 1. TRANSITION PROBABILITIES IN RADAR OPERATOR EYE MOVEMENTS

fixation at t + 1

	Fight	Target	Stranger	Besel	Tote	Console	Other
Fighter	.399	.526	0.0	.045	.007	.003	.021
Target	.358	.492	.019	.073	.005	.034	.019
Stranger	.154	.462	.308	0.0	0.0	.077	0.0
Bezel	.270	.542	0.0	.083	.021	.021	.063
Tote	.231	.077	0.0	0.0	.692	0.0	0.0
Console	.051	.282	.051	.026	0.0	.564	.026
Other	.280	.400	.040	.040	0.0	.040	.200

Figure 2 TRANSITION DIAGRAM SHOWING PROBABILITIES GREATER THAN 0.05



patterns of eye movements are seen during the early stages of practice, but after several hours of practice the requency of fixations become closely to resemble those predicted by the Sampling Theorem. The obvious conclusion is that each meter has, as it were, constructed a model of its forcing function in the observer's nervous system, and the statistics of the model are driving the sampling algorithm. The quantitative properties of such models are certainly not accessible to consciousness, and indeed as implied by Senders findings with pilots, not even the sampling behavior is available to consciousness. But if the sampling behavior matches the requirements of the Sampling Theorem that implies that the statistics are known to the nervous system.

The degree to which the sampling behavior can match the demands of the Sampling Theorem will in turn obviously be determined by the accuracy with which the observer can (unconsciously) estimate the statistics of the time series. At conscious levels it is known that the human operator is not particularly good at statistical estimation. Several reviews of "man as an intuitive statistician" are available, and there is good agreement among them. (See Peterson and Beach, 1964; Dale, 1968; Schrenk, 1969.) Man is quite good at deducing probabilities from frequency of occurrence, and at interpreting probabilities which are not frequencies, ("there is a 0.5 probability of rain today"). On the other hand man is disproportionately influenced by events which occur early in a series, and is unwilling to change on the basis of later evidence. He considers too few hypotheses when more than one can account for the pattern of data. He is bad at estimating variance, being too conservative, and is conservative in combining probabilities. He tends to demand too much information and to make less than optimal use of what he has, and waits too long before coming to a decision in situations involving the sequential processing of information. However, given enough practice, he tends to be surprisingly near the optimum despite these limitations. A similar set of conclusions has recently been offered by Hopf-Weichel, Lucaccini, Saleh, and Freedy (1979) in their review of the way in which pilots deal with emergencies.

In view of the last property mentioned, it is necessary to consider the relation between conscious and unconscious statistical inference. It is abundantly clear that often humans can perform tasks far more efficiently than one would expect from their consciously expressed knowledge of the situation. Consider again Senders et al.'s experiment. In the early stages when the sampling behavior is under conscious control it is relatively inefficient. Given enough experience of the task and the time series generated by the sources the sampling behavior becomes increasingly close to that predicted by the Sampling Theorem, but at the same time the operator becomes less and less aware of what his eyes are doing. It appears that learning, by which information becomes directly able to control behavior without conscious decision making, is an effective way of by-passing the limitations of memory and consciousness. The acquisition of accurate internal models which can guide behavior, in this case sampling behavior, is the requirement for efficient behavior in a dynamic environment. A similar effect will be discussed later with regard to the control of action during plant failures, where there is evidence for the human's ability to compensate unconsciously for changing plant dynamics without being aware of what is happening.

For a well practiced operator, (the word will be used indifferently for a controller, a monitor, or a supervisor,) the control of sampling is usually unconscious, and driven by an internal model of the environment. Only occasionally will conscious control override the model's sampling tactics. The rarity of such events is, however, no measure of their importance, for characteristically such moments are associated with the occurrence of abnormal states or emergencies. Although there is almost no literature on how an abnormal event interrupts the ongoing behavior, when such an interrupt occurs conscious control of behavior supervenes, with its limitations of memory, decision making and information processing, but its compensation in the form of innovation and the possibility of generating actions which have never before been used. The internal model driven behavior can only show well practised behavior.

Under normal conditions, in the absence of unusual states of the observed system, monitoring behavior is limited by two requiremets. On the one hand the sampling tactics must be matched to the state variable dynamics. That is, the sampling mechanisms must be switched from one variable to another at instants determined by the bandwidths of the signals, suitably modified by costs, payoffs, etc., as will be described below. On the other hand, each variable must be examined for long enough for the information needed for an accurate estimate of its value to be made.

There is a potential conflict between sampling rate and sample duration. Although in certain specific cases it may not be apparent, in any large array of instruments it is likely to occur. The nature of the conflict can be seen by considering some data from a study of the eye movements of radar operators by Moray and Richards (1979, unpublished). These are shown as a transition matrix and a transition graph in Table 1 and Figure 2. These data were obtained by analyzing eye movements every second. Thus a "self transition" implies that the operator looked at a feature of the radar display for more than one second, and the distribution of fixation durations can be obtained by raising the probability of a self transition to a power equal to the duration of a fixation. For example, if the probability of a self transition is $\underline{p}$, then the probability that the fixation will be as long as 3 seconds is $\underline{p}^3$, etc. It will be observed that the feature of the display called "Tote" is very seldom fixated, but that when it is, it will tend to be fixated for a very long time ($p_{t-t} \cong 0.25$). This is probably because the amount of information on the Tote board is large, and its legibility is low. If it is to be accurately read it must be fixated for a long time. But to do so means that other features of the display cannot be examined during that period, and important events might occur and be missed.

The transition graph of the radar display is relatively small and simple by comparison with that which would be associated with the instrument panel of a large nuclear power plant, and the mean first passage times for any large control panel will correspondingly be long. It is clear that a first requirement for the efficient design of displays must be to minimize the fixation times needed to extract information, thus freeing the observer to pursue optimal sampling intervals. But what in fact are the latter? We have so far considered only systems where the displays all show equally important information. One of the first criticisms of Senders et al.'s approach was that this assumption is unrealistic when we consider real displays. Different types of information have different value, and they are not in general independent (Carbonnell, 1966; Sheridan, 1970). In particular Sheridan proposed a sampling strategy for a single channel monitoring task in which payoffs played a central roll. We shall now develop some of his ideas further.

Sheridan assumed that a cost was associated with an observation, and that therefore the longer the delay between observations of a variable the smaller the average cost per unit time. But when an observation is made the information remains valuable for some time later, even when the observed variable is stochastic, since bandlimited signals have an autocorrelation function which is nonzero for a time related to the inverse of the bandwidth. As time passes, however, the observation becomes of progressively less value as a means of predicting the current value of the variable, and in the limit, when a very long time has passed, the best guess for the value of the variable is unrelated to the observation, and must be the expeced value of the distribution of the variable. If the object of the operator is to maximize his gain, which is a function of the accuracy of control and the cost of observations, he should take his next sample at the moment when the difference between the value of the last observation and the cost per unit time is at its maximum, since for many common processes this difference function is nonmonotonic and has a well defined peak at some time after the last observation.

Although Sheridan discussed his sampling rule in the context of controlling a single/continuous variable, it is fairly obvious how the idea can be generalized to give a theory of supervisory observations. Consider for example a series of discrete signals, some of which are targets to be detected. These might be radar echoes, spoken words over several loudspeakers, or the kind of meaningless signals used in classical studies of selective attention (Ostry et al., 1976; Moray, Ostry, Fitter, Favreau, & Nagy, 1976). Although in laboratory studies it is common to make trials statistically independent, this is usually not true in the real world. Given an observation of the position of a radar echo future positions are rather constrained by the dynamics of the aircraft. Hence the occurrence of significant information will not in general have a uniform distribution of probability in time, there is an analogy with Sheridan's use of the autocorrelation function for continuous variables, and there will be a distribution which describes the monotonically decreasing value of an observation as time passes.

What are the costs associated with an observation? They may be directly monetary costs. To make an observation may require the expenditure of energay, the hiring of a communication channel, or the use of observers who must be paid. In such cases the costs are explicit and objective. The observer may develop his own ideas about the relative importance of different aspects of the task and formulate them to himself in cost/benefit terms. These are explicit subjective costs. But the most interesting possibility, from the point of view of Sheridan's model, is to approach the problem of cost from a non-monetary point of view, defining costs in terms of the structure of the task, that is the probability that observing one variable will cause important information from another source to be missed. The observer may define the cost of an observation as the probability that while making it he may miss the moment when an observation should be made on another channel, that the error on another channel may exceed the allowed error, or that a target may occur on some other channel while he is sampling the chosen channel.

The approach is closely related to queuing theory and time line analysis. Let us suppose that the basic sampling rule is to sample at T_S , where

$$T_S = \max(V_t - C/T)$$

and V_t is the value at time of an observation made at T_0 , C is the cost of an observation, and T is the time since T_0 . For each source there will be a similar equation defining its sampling interval, so that we have a set of equations

$${}_iT_s = \max({}_iV_T - {}_iC/ {}_iT)$$

where i-sub denotes the ith source.

Now C is the "selfcost" of an observation, that is the cost of the observation if no other source were present. But if there are two channels we must weigh the cost of each by the probability that the other task will have a sampling interval at the same time. If there are more than two channels the argument is extended to include weightings for all other channels, giving as a first approximation

$${}_iT = \max({}_iV_T - ({}_iC/ {}_iT + {}_{j\neq i}P_j)$$

where P_j is the probability that source j has its next sampling instant at time ${}_iT$. Note that P_j will be a function of several variables, including bandwidth, permissible error, payoff etc., of source $\bar{j}$, and that the weighting for value for $\underline{j}$ itself is included in its calculation.

Sheridan and Rouse (1971) and Rouse (1972) found that human operators were suboptimal in terms of Sheridan's model, due to using less than the optimal amount of information for prediction. It is more surprising that laboratory experiments can produce behavior even fairly close to the optimum, for as we noticed when discussing Senders' experiment, to do so requires that the operator has performed the task long enough to acquire good knowledge, probably at a subconscious level of the statistical structure of the task, the implicit and explicit payoffs, etc. Indeed the amount of information to be processed by an operator in this kind of task is formidable; all the more so when one considers the limited rate at which conscious decisions are made and that many sources may not be statistically stationary in the real world.

Costs could be made explicit during training, provided that care were taken to ensure that the operator both paid lip and behavioral service to them, and not merely the former. (Various authors have reported from time to time that the verbally expressed behavior and actual behavior may be quite at variance with one another.) But the relative probabilities of signals, the probabilities of simultaneous demands, etc., will all be estimated by the human operator under the constraints of his qualities as intuitive statistician, which we noted above. In particular, rare events, and very low bandwidths are likely to prove of great difficulty for the operator. If in fact he overestimates their probability and frequency, this will be to his advantage, in so far as it will ensure that they are sampled at least often enough, although too frequent sampling will reduce the frequency with which other sources can be sampled. The greater danger is that low frequencies and probabilities will be treated as if they are zero, and such sources will, as practice progresses, become effectively locked out altogether from the sampling schedule. It should be noted that classical learning theory would predict this, in so far as an observing response which is never reinforced (perhaps by detecting something of importance) will extinguish and disappear from the behavioral repertoire. There is a clear need for research into the question of how to train operators to handle very rare events. It may be that the application of operant conditioning rules are applicable. It is known that providing appropriate training schedules are used it is possible to produce behavior which persists for a very long time with little or no reinforcement. But the point is that such is only the case where the behavior is carefully shaped during the acquisition phase. Merely letting the operator sample as he will is most likely to result in the disappearance of rare events from the sampling schedule. (If we translate this from sensory data acquisition to the level of cognitive hypothesis testing, it may account for the tendency of the human operator to test too few hypotheses when interpreting data. The maximum flexibility in hypothesis generation must be trained; it will not occur spontaneously.)

It seems likely that one approach to the monitoring of rare events in supervisory control may be through the use of new computer based technology. Given the relatively trivial cost of computing, it should be possible to supplement the humans' sampling with a computer prompting system. We shall discuss later the rules which such a system should use. At the moment it will be sufficient to note that with conventional instrumentation it is not usually possible to note which instruments the human operator examines, since the recording of eye movements is not practical in a complex environment in which the operator can move freely. On the other hand, if extensive use is made of VDU's, and the operator has to call up the particular display which he wishes to sample, then the computer could keep a log of which sources have been examined, and could prompt the operator from time to time if some variables had not been examined for a considerable period. Indeed in principle the computer could go further, and model the distribution of sampling shown by an individual operator, model the bandwidth of the sources, and offer advice designed to help the operator converge on an optimal sampling strategy. There has been very little use so far of computer systems which adapt to human users, but it is now economically feasible to add such facilities to computers which control automatic processes, in order to make use of symbiotic control between the man and the computer. There may be difficulty in finding the optimal level of symbiosis, such that the operator appreciates the assistance of the computer without coming to rely on it exclusively.

Speed-Accuracy Tradeoff

The discussion so far has made an implicit assumption, namely that the fixation time, dwell time or (in queuing theory terminology) the service time is short compared with the sampling interval, so that there is no conflict between the duration of data acquisition and the need to switch the sampling system to a new source. That is, we have assumed that the service time is negligible compared with inverse of the bandwidth of the system. That assumption is not in general valid, and brings us to the second source of difficulty ofr the supervisory operator. If the service time were indeed negligible, then we might already have the main outlines of a model of supervisory control, and the main problem would be to ensure thta training was adequate to ensure the correct estimate of optimal sampling instants. Moreover, even if demands to sample came from more than one source at the same instant no significant loss of information would result, since the negligible service time would allow first one and then the other source to be sampled. The only real source of interference would be on the response side. But the time taken to acquire data from a source is not negligible, as we have seen from the example of the radar operator.

We noted that Senders et al. (1966) originally suggested that service time would be proportional to the information content of the source. It is widely accepted that information content usually does cause the time to process signals to vary, although there is some controversy as to why (Attneave, 1953; Garner, 1962; Kornblum, 1975). In general we may expect that in a multiple variable system there will be considerable variations in service time, since the bandwidths of the sources will differ, the signal to noise ratios of the signals will differ, and the permissible errors will differ, all of which will cause the effective information content to differ in turn. In the laboratory it is possible to produce situations in which the performance of the human operator is largely independent of information content (Mowbray and Rhoades, 1959; Leonard, 1959; Davis, Moray and Treisman, 1961) but the necessary conditions are unlikely to occur in industrial settings. In general, the more difficult a signal is to discriminate, and the less probable it is, the longer a human operator will observe it before making a decision. Human operators behave as if they accumulate information while they observe a signal, and wait until some subjective criterion is satisfied before they make a decision.

There is considerable evidence that a good model for the human observer is a sequential decision maker. At extremely short durations both the auditory and visual systems behave as though they were integrating energy at the receptors. Over periods of a few hundred milliseconds the longer a stimulus is presented the more detectable it becomes. For longer signals and situations more relevant to real applications the best evidence comes from reaction time studies (Laming, 1973; Vickers, 1970; Pachella, 1974; Green and Luce, 1973). For example, in a number of studies it has been found that the index of detectability from the Theory of Signal Detection (TSD), namely d", increases almost linearly with the response latency. Taylor (19) has argued in a more general framework that information is accumulated proportionally to the square root of the signal duration. Moray (1979c) has also shown that the TSD response bias criterion, beta, increases with response latency. This is almost a trivial result, since obviously if a very fast response is required the minimal evidence must be used with a high attendant risk of false alarms. But Moray also noted what appeared to be differences in strategy between his subjects. If there were a constraint on the time available for making a decision so that the response had to be made before a deadline, the criterion first rose, then remained steady at its statistically optimal value, and then just before the deadline showed huge individual differences. One operator showed a large rise in beta at the deadline, another a marked fall. These results are what one would expect if initially the operators set appropriate criteria ofor the properties of the signals, but that faced with insufficient evidence and a deadline, one decided never to say "yes" if in doubt, and the other never to say "no".

The relation between the detectability of a signal, the response criterion, and the service time, is very complex. Two demands must be satisfied, the scheduling algorithm which tells the operator which source of information to sample next and the time at which he should take the sample, and the data acquisition procedure which decides when enough data have been received for an accurate decision to be made about the state of the variable being sampled. If we assume that the operator uses some kind of sequential decision procedure for the latter purpose, for example the procedure suggested by Wald, (1947) then he must make use of his instructions, experience, and his understanding of the payoffs associated with the task to choose suitable response criteria and set the decision boundaries appropriately. He will also be preset to sample another source at some time in the near future, due to the scheduling algorithm. As he examines the source of information he will begin to accumulate evidence aobut the value of the variable. One of two things will next happen. He may gain enough information to make an adequate decision about the variable, and thus be freed to switch to another source; or the urgency of sampling another source will become so great that he will be forced to make a decision on the basis of inadequate evidence, (for example, by assuming that the correct decision is the nearest boundary in the Wald space,). Alternatively he may have to delay his switching to another source and thus violate the demands of his statistical models with regard to sampling. Thus any display with degraded information will tend to cause either suboptimal decisions, or interfere with sampling. Research is needed on optimising the way in which information is displayed on VDU's since they are usually digital displays. For example, Hess and Teichgraber (1974) report that if display information is too coarsely quantisiseded in a digital display it increases task difficulty as measured by a secondary critical task.

This conflict between the need to obey the sampling schedule and the need to acquire adequtate data will tend to propogate error throughout the whole pattern of supervisory control. If the operator curtails his examination of the variable in order to switch, he will have a less than optimal estimate of the state of the variable, which may result both in his missing a state which calls for some action on his part, and also missing information which would lead to a change in the sampling schedule. On the other hand, if he does not sample the next variable to which the sampling schedule is calling him, he will lose information about that variable with the same effect, and may have a very disruptive effect on the whole schedule. The concept of speed-accuracy tradeoff was mentioned earlier. Here it may be given an extended meaning. It has been traditionally taken to refer to behavior on a single trial, but if applied to a pattern of complex behavior it is clear that trading speed for accuracy or vice versa may not be detrimental. To accept errors on a particular channel may be preferable to corrupting the long term sampling behavior which guarantees that the entire system is adequately scanned.

Recently Tulga (1978) has found direct evidence for speed-accuracy effects in a multi-task supervisory situation. His observers monitored several queues of signals with different values and durations, and were required to schedule service to the queues so as to maximize their reward. He found that at low processing rates operators chose a task, not a schedule, since there was time to perform each task in turn, and a schedule was of no advantage. Subjectively a schedule was experienced as a harder task. As the load increased the operators at any moment tended to discount queues which has slack time and to concentrate on busy queues. Furthermore they began to take less account of future states of the system which they could stochastically predit from the display. Tulga also measured subjective workload and found that when the imposed workload was so great that the operator was overloaded the subjective workload actually decreased, but so did performance. One must conclude that the operator was adopting a less stringent criterion of accuracy in the face of a situation demanding the maximum rate of work.

It may be that there are significant implications here for training operators of multi-degree-of-freedom systems. If there are many variables it is almost certain that there will not be time to observe each source for as long as the operator wishes, and training should emphasize the necessary tradeoff. In the limit, if the operator decides to stay on a particular channel for a very long time because he cannot interpret the information he is receiving, he will, of course, completely lock out all other information. Accident research suggests that such events are not uncommon, and we shall return to the point when discussing error diagnosis. Here we may note that new kinds of displays may, if they require the operator to call up information with a keyboard or other control device, acutally increase the difficulty of the task if the time taken for the system to respond (including the time to enter the commands) is longer than it took him with a conventional system to move his eyes or walk from one part of the display to another. It is clearly of the greatest importance that computer controlled displays be engineered in a way which minimizes data access and acquisition time. A minimal design criterion should be that less time is taken with new displays than was taken with conventional displays for equivalent accuracy.

There is some evidence that sampling strategies do indeed change under stress. Cedr (1977) found that as traffic became more congested the eye movements of automobile drivers became smaller and fixation time increased. That is, as the information rate increased, less of the environment was examined, and examined for longer. On the other hand Clement, Hofmann, & Graham (1973) found that pilots in a simulator showed fixations which were of almost constant duration at around 0.4 secs. (They do not describe sampling behavior.) The paper by Moray et al. (1973) already cited suggests that where the rate of information is under the control of the operator the rate may be adjusted to keep sampling rate constant. Sperandia (1971, 1978) found that as the number of aircraft increased air traffic controllers reduced the number of variables per aircraft which they monitored. When the total load became too high, they either called in a colleague, refused to handle the aircraft, or constrained its behavior into a rigid flight pattern which reduced the amount of information which they had to handle per unit time. It is clear the latter strategy cannot often be used.

Nor is it obvious that to constrain the task to a rate of information preferred by the operator is necessarily efficient.

The design of appropriate displays for supervisory control must take into account the interplay between the two kinds of sampling decisions. Correct scheduling of sampling may require computer aiding, and display clarity must be such as to reduce data acquisition time to a minimum. While classical queuing theory modified by taking account of the relative importance of the sources of information and the costs associated with observations may suffice to settle the scheduling problem, we must still consider the problem of optimizing data acquisition, since as we saw earlier man tends to wait too long and demand more information than is really needed for a decision.

It is abundantly clear that the factors which influence the way in which he acauires data are many and complex. For example Levine and Samet (1973), and Levine, Samet * Brahler (1975) report that operators are sensitive to the quality as well as ot the quantity of information. If the reliability of the data fell, more information was requested before a decision was made. High reliability improved both speed and accuracy. As the conflict between sources due to time pressure increased less information was acquired from each. Sources whose information was of high diagnostic value were preferred over other sources. These results are in good agreement with the general model sketched above. They suggest that a computer might be used ot track the diagnosticity of the information which it displays by logging which sources are used by the operator, and then might be used as an adaptive trainer to ensure that on the one hand the more diagnostic sources were emphasized, and on the other that all sources were at least occasionally sampled.

At the theoretical level there appears to be good agreement about the optimal amount of data needed for a decision. Mullis and Roberts (1968) show that in sequential decision making 8 samples are almost as good as an infinite memory for determining the detectability of a signal in noise by taking successive observations. Baxa and Nolte (1972) also claim that 8 samples is optimal. Taylor (1975) investigated the best way to discount the past so as to achieve an optimal running estimate and found that with a memory of less than eight samples the quality of the decision rapidly fell, but that above 8 samples the gain in quality was negligible. (It is interesting to observe that, except under very exceptional conditions, the immediate memory span of humans is about 8 items, and to speculate as to whether this is some kind of evolutionary discovery of the optimal memory size for making the kind of conscious decisions required for adaptive behavior in the real world.)

Given this agreement about the theoretically optimal size of memory for decision making it may be possible to derive an analytic model for the interacion between the amount of information taken once a source is selected, the frequency with which a source is selected, and the overall efficiency of an observer in monitoring multiple sources of information. The problem of validating such a model of behavior is, however, formidable. When sampling rates are slow, as in the experiments of Senders et al. one is inclined to assume that each time an instrument is fixated a single sample is taken, (although Senders did make provision for multiple looks without a change of fixation). On the other hand, where sequential decision theory has been applied to the analysis of, say, reaction time experiments, it is usually assumed that successive observations take place within a time scale of milliseconds and that there is no way of observing their occurrence. Whatever the solution of this problem, it suggests that displays might be improved by considering how to choose the optimal amount of data in the light of these theoretical studies. It would presumably aid efficiency if an operator could be trained never to take more than "8 samples" from a source, however we may interpret that instruction. One attempt to develop such an approach was made by Clement, Jex and Graham (1968), who used it to make recommendations as to the optimal design and layout of instruments for the cockpit of a jet transport. Their paper represents the most detailed application of such ideas available in archival journals.

In general human operators are conservative in their decision making, but there are exceptions. Kvalseth (1978a,b) reports a study in which observers could choose the amount of information they wanted, and used only one observation where 4 observations would have been optimal. But there is a basic problem in interpreting all studies which purport to show that the human operator "uses only N samples", "tends to be conservative", etc. Such results are no doubt true <u>for the conditions of the experiments where they were obtained</u>. But the human operator is neither deterministic nor, really, stochastic. He is strategic, and his behavior is a reflection of the way in which he views the task. We may recall the remarks by Rasmussen which were quoted at the beginning of this paper, concerning the richness of strategies available to operators in real life tasks. The provision of high quality displays is a necessary but not sufficient condition for efficient supervisory control. Another necessity is to ensure that the strategies which he uses are appropriate to the structure of the data sources. Man is a strategic processor, not a transmission channel or linear control system (Tamura and Yoshida, 1974). The complexity of modelling the operator is emphasized by a study by Monty (1973) on the ability of operators to keep track of several variables. He found that all the following were relevant: the rate of presentation, the number of categories, the length of the run, the ontime-offtime ratio of the displays, whether or not a signal was cued, the regularity of presentation, the occurrence of more than one variable in the same position on the display, and the presence of secondary tasks. (The fact that when more than one variable occurred in the same position performance was less efficient should be considered in the light of the change to general purpose VDU's) Probably the simplest assumption that can be made is that an operator will assume that every aspect of the display is relevant until he has been trained otherwise.

When a display is sampled, what is it that the operator is looking for? In most supervisory control it is to see whether the values of variables are within an allowed tolerance about the set point. If such is not the case then the operator should intervene. We shall return to this point later when discussing error detection and fault diagnosis. But the great part of the evidence suggests that in monitoring automatic systems such as aircraft, process control plants, and such like systems, the human operator is examining the magnitude, sign and rate of error signals.

A final aspect of information acquisition is individual differences. Since payoffs and probabilities have to be estimated, there is scope for individual differences to play a considerable role in data acquisition strategies. For example, there is a persistent hint in the literature that there may be important differences related to "field dependence". This dimension of personality has been found to have small effects in a number of studies, and relates to the relative tendency to depend on sensory information or to be more analytic and independent. Norman (personal communication) has referred to the difference as being between "data driven" and "model driven". If such differences are reliable they may be central to differences of personal styles in supervisory control. Goodenough (1976) found that field dependent automobile drivers used only visual information, and did not combine it with information from other sense modalities. They were slower to detect abnormal situations, slower to respond to the movements of another car in a following situation, and appeared to make little use of distant cues, concentrating rather on the small area of the visual field immediately in front of them. By contrast the field-independent drivers made better use of peripherally presented information and were able to synthesis cross-modal information. Robinson and Bennick (1978) present related results, and Ragan et al. (1979) an extensive literature review of cognitive style as reflected in field dependence.

If such findings are general they suggest that the field dependent operator is someone who will tend to become fixated on a relatively small subset of the available displays, and who will show less ability to synthesis information across displays and make use of his general knowledge and his internal models in scheduling his behavior. This would mean that field dependent operators would find it harder to devlop efficient sampling schedules, and would tend to give them up in favor of data acquisition under stress. If the effect is large, there would be a strong argument for trying to select field independent people for training as supervisory controllers of complex systems, or for looking for ways in which the computer could interact with the operator so as to offset any tendency towards field dependence. It would certainly seem worthwhile to screen operators for field dependence when very large and complex systems such as nuclear power plants are involved.

The Incorporation of Information

There are a number of different ways in which humans store information to which they have been exposed in the past. There is, for example, a real difference in the way in which the words "memory" and learning" are used, and even the work "storage" would be somewhat misleading, since it implies the holding of an item of information in a store, and does not convey the flavor of enduring changes of organization such as learning a skill or a transfer function require.

Memory, Learning and Models

It may be useful to restrict the word "memory" to refer to the kind of information which the operator can access voluntarily and express verbally, or at least symblically. Thus a set of instructions would be stored as reportable contents of (long term) memory, while the knowledge of how to ride a bicycle or whistle would not. In the latter two cases we will say that the skill has been learned and retained. In some cases the results of learning can be described verbally. Indeed if such were not the case it is hard to see how athletics coaching or training in general could be possible. But there are many perceptual motor skills which are possesed by humans but which cannot be translated into verbal accounts, but which are efficient, durable, and complex.

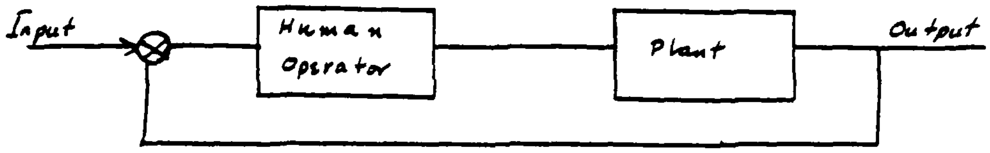
The concept of a "model" can perhaps mediate between two sorts of information incorporation. In this paper a "model" is an internal representation of the statistical structure of the world, sufficiently accurate to allow an operator to predict future states of the world from a small sample of information currently received through the sense organs. It may include both verbalisable and unverbalisable features. It may include movements, skills, strategies and tactics. It may be altered and upgraded by new data and by internal reflection upon those of its features which are accessible to consciousness. It is constantly changing under the impact of new information. Its most important property is that is is predictive. A human operator possesses an array of models for different environments and different tasks, and one problem for him is to choose the model which is most appropriate for the particular task with which he is confronted. There is some evidence to suggest that not more than one model at a time can be used, although some may be hierarchically nested inside others, rather like subroutines in a program. A stimulating account of models in this sense will be found in Kelly (1968).

A central problem in model acquisition and use is the efficiency with which information can be held in short term memory. There is broad general agreement that such information will be lost if it does not enter long term memory within a few seconds, but that information which once enters long term memory is never lost, but may become inaccessible, and suffer from interference from similar information during retrieval. Little direct work is available on the short term retention of skills, and on the kind of problems involved in supervisory control.

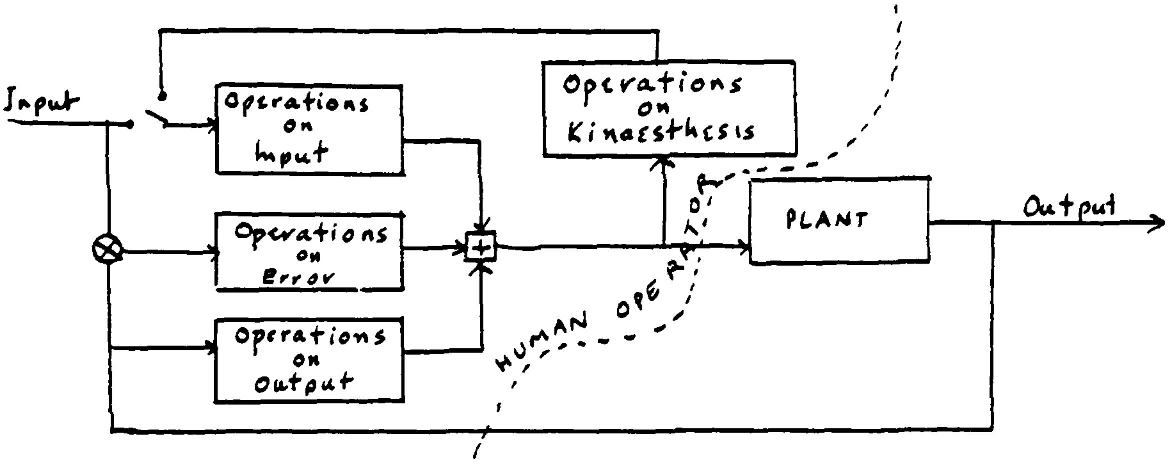
The most important problem in supervisory control which involves memory would seem to be "book keeping". Given that a large number of displays must be sampled, whether they are all physically present or must be called up on a VDU, the operator must keep some kind of a log, either mentally or physically of the order in which he has sampled the variables, or at least the length of time which has passed since each source was last sampled. Prolonged practice leads to the operator acquiring a model of the process which embodies the statistical structure of the information in time and space. The solution of the sampling problem is a schedule which moves the data acquisition system among the sources in a way which can be summarized sometimes as a Markov process, and which is driven, it seems likely, by such properties as the bandwidth of the sources, modulated by their importance and the payoffs associated with each (Sheridan, 1970).

The human operator is seldom if ever aware of the schedule which he adopts. He may be aware of some gross features of the model, but the detailed model is not accessible to consciousness. An attractive picture of how models are acquired is given by Krendel and McRuer (1960), reproduced in Fig. 3. Although their description is in terms of a single channel closed loop tracking task the ideas may be generalized in a way similar to the generalization of Sheridan's Supervisor Theory. The critical property of their picture is the emergence of "programs" at the third stage of skill acquisition. In order to be an efficient supervisor one or more of these programs must contain the information which allows the calculation of priorities to establish which source should be sampled next. The model will then drive the sampling behavior independently of any values of data found during sampling. We might, to be specific, assume that the programs were each a Sheridan Supervisor dedicated to a particular data channel. Each would then bid for the use of the data input channel, (for the right ot determine the direction of visual fixation, or the right ot call up its data on the VDU,) on the basis of its current distance from the maximum of its cost/value function. If a clash occurs due to equally urgent calls from two or more Supervisors at the same moment, then some kind of random choice among the contendrs will be made. (We assume that if there were any secondary information available which would help to arbitrate among the contenders it would already be incorporated in the weighting function. Hence a clash of interest by definition leaves only random allocation as a means of resolving the conflict.)

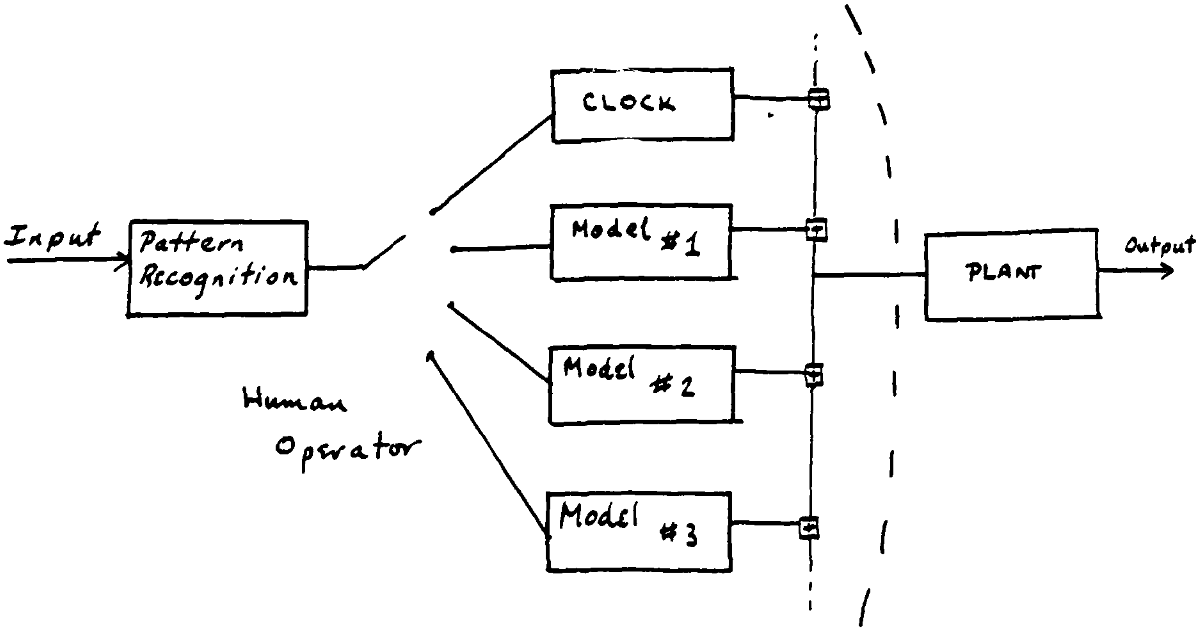
The basic problem for memory and learning then is to acquire the necessary information during practice to build models, and to update them from time to time on the basis of experience. We therefore wish to know what kind of things can be modelled and how efficiently the models represent the world. Note that Krendel and McRuer assert that not only sensory information, but also outputs and control procedures themselves can be modelled.



a. Compensation



b. Pursuit



C. Predictive Open Loop

Figure 3 (After Krendel & McRuer)

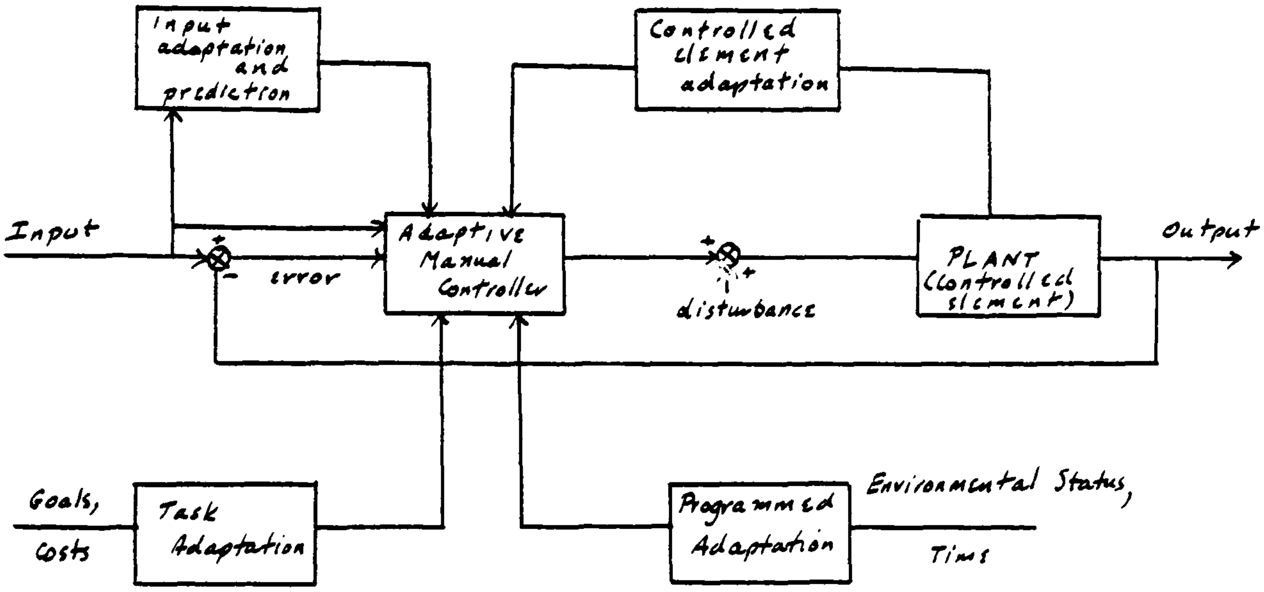


Figure 4 : Young's Adaptive Controller

The operator may acquire "motor programs" as they are currently called, (see e.g. Keele, 1970; Bernstein, 1967), or models of the relation between his actions and the effects they produce on the controlled element. In the latter case the operator comes to have at his disposal a set of transfer functions among which he can choose in order to equalize the properties of the controlled element (McRuer and Krendel, 1974). Features which can be incorporated into the internal models of the adaptive human controller are described by Young (1969) (Figure 4).

Short Term Memory

The severe limitations on human short term memory renders the task of supervisory control very difficult. Memory is needed to accumulate evidence about a single source, and even more to combine evidence from several sources for the purposes of diagnosis. Memory is needed also to update the internal model, and to make sure that the sampling schedule is carried out. If the nature of the sample results in a disruption of the sampling schedule, or if some emergency arises which distorts the normal pattern of sampling behavior in order to deal with a particular source or sources, then memory is needed to re-enter the sampling sequence at the correct place. Otherwise there may be extremely long intervals between samples from some sources, with a resulting high chance of abnormal states being missed if they occur.

We have seen that in theory a 3-bit (8 item) memory should be adequate for sequential decision making. But there are several reasons why the human supervisor will not be so efficient. The first is that man's dynamic short term memory is much smaller than 8 items. The laboratory evidence very strongly suggests that running memory is not more than 3 items, well below the optimal length. In a sequential probability task Baker (1963) found that his subjects only used the last three intervals to predict the next. Michaels and Moray (1979, unpublished) found that observers seemed only to use the last interval in a series when estimating future inter-stimulus intervals. Seitlin and Finkleman (1975), Meckworth (1959) and Kay (unpublished) have all found memory spans of three or less in a task which has face validity as being relevant to aspects of supervisory control. The participants watched a series of lights turn on and off were required to indicate either the light currently on, or the previous light, or the light that had come on two intervals ago but was not no longer lit. All these workers agree that the last task is almost impossible for the human operator. Moreover Zeitlin and Findleman found that the two-back condition imposed a very great load on the operator when he was required to perform a tracking task at the same time. We may conclude that holding even very simple information in memory for subsequent use, while simultaneously entering new information and removing old information extremely difficult for humans.

Short term memory is further degraded by any cognitive tasks which are performed while holding material in memory. The "Brown-Peterson Paradigm", in while a subject is required to count backwards while retaining material in memory has for years been used as a classical method for preventing the permanent storage of information. Since in real life control tasks almost everything about the task except the currently observed variable would constitute an interfering task, the control and observation of complex systems is likely to tax memory to the uttermost. Both Loftus (1979) and Moray and Richards (1979, unpublished) have found rapid forgetting for "real" material. Loftus found a 15% loss of information in 15 seconds for status information among Air Traffic Controllers, and Moray and Richards found that memory for simple simulated radar displays began to fade rapidly after about 12 seconds. Loftus makes the point that traditionally ergonomics has concentrated upon the efficient presentation of information, and that perhaps as much attention needs to be paid to efficient ways of allowing the operator to off load information from memory before it becomes inaccessible.

These limits on short term memory may account for some of the difficulties operators have in performing sequential decision tasks. We have noted the limits of man as an intuitive statistician. Rouse (1978) found that in fault diagnosis in a logic network operators were good at using the topology of the net to make initial diagnoses about the position of a fault, but poor at changing their opinions in the light of later evidence, even with a display in front of them. If in fact their thinking were erasing the memory for the information which they had already stored, this would not be surprising.

In general we may conclude that the task of monitoring a large display with many instruments is one for which the human memory is ill suited, especially when it is necessary to combine information from different parts of the display and the information in dynamic. Aids to efficient memory are badly needed.

In certain respects the rapid loss of information might, if the task could be organized properly, be a good thing. It is in fact rather curious that the human operator is so conservative and so heavily influenced by events early in a sequence if he tends to forget so rapidly, since this should tend to abolish the effects of early items in a sequence. It is known (Gonzales and Hovington, 1977) that discounting earlier parts of a sequence and using only the last few items for estimating the state of a process can be used in a very efficient method of detecting errors, and the rapid memory loss in humans should predispose them to adopt that kind of approach. One must assume that the extremely short span of running memory is just too short to be effective. (Gonzales and Hovington, op. cit. describe an automatic error detection algorithm which is more efficient than the human, and can detect errors as small as 0.1% in process variables at Oak Ridge.)

It is a paradox that given a poor memory the human operator is unwilling to change his mind in the face of evidence (Schrenck, 1969; Dale, 1968). Recent work has shown the same effect in prediction extrapolation. For example, Taylor (1975) and Gai & Curry (1978) found a marked conservatism based on the first few elements in the sequence displayed.

It seems then that memory is both vital for efficient monitoring and inefficient in two ways, biassed towards early events and limited in size. In view of this, how do human operators become skilled, let alone reach the striking levels of efficiency which they frequently show?

Internal Models

The human operator knows that his memory is limited and when possible will take steps to reduce the load (Sperandio, 1961, 1968). The bandwidth of many industrial processes is very low, and errors which the operator makes are usually noticed and corrected before they have time to affect the system performance. In this respect the report by Ruffle Smith (1979) on flight crew errors is particularly interesting. On the average, errors occurred as often as once every five minutes. Some of them were due to forgetting what happened earlier, others by forgetting the readings of instruments or the values of variables which were needed for decisions. The complex tasks of real life environments show memory at its worst, and conscious memory will suffer from severe mutual interference between different subtasks unless artificial aiding is provided.

On the other hand information about the task may be incorporated into the operator's internal model, of the contents of which he is unaware. Indeed a conscious description of his behavior by an experienced operator is frequently very inaccurate. Senders' subjects might have been aware of the rank order of the bandwidths of the instrumets which they observed, but they could certainly not have described the transition matrix which was necessary to model their eye movements. Yet their behavior was tightly coupled ot the statistical properties of the display, and this coupling we have ascribed to their models.

How then are models and memory related? Models are learned, not remembered. The information on which they are based may for short periods be accssible to consciousness, although not always, and not necessarily. For example, Pew (1974) describes an experiment in which the middle portion of an otherwise random forcing function was repeated on each trial. The operator became increasingly efficient at tracking the portion that repeated, but was not aware of what was happening, even though obviously aware of the moment to moment value of his error. If we express the situation in terms of Young's diagram (p. above) the operator will be aware of his instructions, of the current value of the input if he observes it, and usually of the environmental status. But he is not aware of the form in which these are being incorporated into the adaptive controller, nor explicitly aware of the control law of the controlled element.

It is unfortunate that there is virtually no research on the acquisition of models. There are plenty of examples of well developed models, both classical control theory and optimal control theory (McRuer and Krendel, 1974; Levison, 1979; Stassen and Veldhuyzen, 1977). Young's paper deals with two kinds of adaptation, the tuning of parameters or the distribution of one model for another when the controlled element characteristics change. Jagascinski and Miller (1978) suggest how one might estimate how accurately a model had developed for a very restricted range of tasks, by tracking the operator's performance on a root locus plot. But there appears to be no study of the course of acquisition of a model, with a description of how the model is built, and the variables which affect its acquisition. One must suppose that learning theory would account for the growth of a model, although there may be problems insofar as learning theory has been developed exclusively to account for the development of behavior, not the development of a controller of behavior. We are here discussing something that is almost the opposite of classical stimulus-response learning, namely a system which will produce a response because it believes that a certain state vector value is about to occur at some moment in the near future. What is required is a system which abstracts patterns from incoming data, and from its own responses, and from these can construct a picture of itself and its relations to a large state vector with which it is interacting.

Some account along the lines of Krendel and McRuer' Successive Organization of Perception seems promising, and the central role of an internal model in processing is worth emphasizing. Indeed Conant and Ashby (1970) offer a formal proof that any efficient controller of a system must logically embody a model of the system in the ocntroler; and while their claim has been contested (Kickert, Bertrand, and Praagman 1978) on the ground that it is only concerned with error reduction and not with guaranteeing stability, their argument gives added force to the need for an adequate account of model acquisition.

Several other workers have pointed to the importance of models, Bainbridge (1978) notes the applicability of the notion to industrial skills. Phatak and Bekey (1969) pointed out that a system to detect failures must have a model of the pre-failure system behavior against which to compare the observed behavior, a failure detection device, the ability to adapt to post failure dynamics, and if successful in controlling the failure a model of the postfailure steady state system behavior. Wickens and Kessel (1979) speak of the operator having a model of the state variables and a channel to estimate the current value of those variables from noisy observations. Moriarty (1974) noted that operators might adjust to a change in controlled element characteristics as they altered from pure inertia to camped second order without being aware of the change until the process was 30-40% complete, and suggested that they unconsciously adopted a series of models during the transition. Wickens and

Kessel also discuss the unconscious compensation for plant failure and the relative role of kinaesthetic and visual information processing.

The fact that such unconscious compensation can happen, suggests two things. Firstly that operators who thoroughly learned a skill retain models which can be called up and used when required, (Krendel and McRuer's "programs",) and secondly that the calling up of models does not require consciousness. Note also that the initial acquisition of a model may require many hours of practice for a skill to develop, but that once the model is available adaptation to changes in plant dynamics may occur in a few seconds.

To validate the existence of models it is not enough to show that the human operator shows adaptive behavior. The time scale involved in adaptation and acquisition of skills are quite different. They can indeed differ by many orders of magnitude. Young (1969) describes adaptation to plant reversal in a few seconds. Pilot behavior in regaining control of an aircraft following the failure of an automatic system may take tens of seconds. Regaining control of the Three Mile Island nuclear reactor took many hours. And Crossman's (1959) famous study of the skill of cigar rolling showed the continuing effects of practice over 7 years and several million cycles of the operation!

The application of classical control theory to human performance inherently is a model orientated approach. The fact that the human operator posseses a set of transer functions and can choose which to use so as to make the overall man-machine system have a transfer function as near as possible to a first order system with a delay and variable gain, (McRuer and Krendel, 1974) is equivalent to saying that he has a set of models in our sense. As mentioned before, the choice of the appropriate functon is efficiently made, but the human operator is certainly not aware of the function he chooses. Behavior is goal orientated, not merely stimulus determined, and these are the basic qualities of models in human skill.

In recent years the emphasis has changed to the use of Optimal Control Theory, and here the emphasis on models as central to skill is even more direct. Also OCT emphasises the fact that learning is not asymptotic but dynamic, and constantly changes the properties of the learner in relation to the environment. Models are continually modified on the basis of incoming information and the effects of action. The heart of current OCT is the Kalman filter which performs real time modelling of the statistics of the state variables, and uses that model to choose an optimal control acton. While the major thrust of OCT so far has been in control, the Kalman Filter is essentially an estimate, not a controller of the system, and recently has been used in the context of error detection, appropriate to supervisory control, (Curry and Gai, 1977; Murthy, 1975; Wickens and Kessel, 1979). The optimal estimator is used to track the supervised process, and a sequential decision algorithm is used to discover when the process has gone out of the permitted tolerance range.

It is now clear that in many cases the OCT approach is very successful at modelling the output of the human controller (Stassen and Veldhuyzen, 1977; Baron, Kleinman & Levison, 1970). It is also, in the context of the present discussion, particularly interesting that Baron and Berliner (1977) showed that a correct model of the controlled process is perhaps the most important possession a controller can have. Neither excessive motor noise, nor an incorrect assumption about the bandwidth of the controlled process made much difference to the performance of the optimal controller. But an incorrect internal model of the controlled process had a very marked effect, causing large errors. If OCT continues to be successful in predicting human behavior, this suggests that training should perhaps be concentrated on helping the operator to acquire good models of the controlled process, rather than directly on producing correct behavior. In so far as supervisory control does not require much overt behavior from the operator, training for models may be a task of considerable difficulty, and one which requires research. It certainly implies that good mimic diagrams of complex systems, which show causality in the system in a particularly direct manner, should be regarded as of the greatest importance.

Research is also needed into how to tell that an operator has or has not acquired a model, especially where there is little or no overt behavior.

Prediction and Extrapolation

Why are models so important to efficient complex skilled behavior, and why has the concept not received more interest from psychologists? The answer to the second question is in fact simple. As Rasmussen (1979) and Moray (in Underwood, 1979) have noted, laboratory tasks are explicitly designed in most cases so as to render to the use of models inappropriate. Trials are almost always independent of one another so as to aid statistical analysis. When trials, or events in a natural process are not independent, then the importance of models becomes evident: they allow extrapolation and prediction. Incoming information can be matched against a model of the world, and a successful match allows the operator to predict in advance what the world will shortly be like. Hence action can be planned and initiated in advance of its being required, reaction time lags reduced or abolished, and the operator freed to some degree from time pressure, since he no longer need process incoming information in order to decide what to do. Again, if sampling is to be optimized a model is an absolute necessity so that the waxing and waning of event probability can be mapped onto real time. Models are the source of strategies and tactics in behavior, and free the human operator from the constraints of deterministic causality imposed by the environment.

It is for this reason that the idea of a model is becoming increasingly popular even in areas where the rigorous quantitative treatment on control theory cannot be applied. For example, Grant (1971) discussing the problems of alertness in long distance train drivers points out that the problem is to combine long term memory of the quality of the track with the need to assume that it will not always be in exactly the same condition. Long term memory serves as a model to generate expectancies as to the places where variations are most likely to be needed, while short term memory and moment to moment information processing takes care of variations specific to the particular run. It is now common to find workers in the area of Air Traffic Control (ATC) speaking of the model which the ATC uses, models of the state of the airspace, or of the overall behavior of the system. Thus Coeterier (1971) describes ATC's as having models of the different aircraft and changing these models as time passes updating them with information from the displays. Moray (1979 unpublished) has adopted a similar approach in modelling the behavior of fighter controllers guiding interceptions by radar. Coeterier argues explicitly that ATC's use their models to predict future movements of the aircraft. Broadbent (Personal communication) has investigated the differences between the model of a process of which the operator is conscious and the actual behavior when controlling an economic model, and has found that the unconscious model tends to be more efficient. Bainbridge (1978) discusses the way in which models are used by process control operators to predict the future states of the system. The emphasis of all these approaches is that information is not simply used as a stimulus for action, but to provide a data base which can be manipulated for cognitive decision making, in line with the modern emphasis on man as a processor of information rather than an S-R system.

Referring back to Young's and Krendel and McRuer's diagrams, we can summarize the ways in which models can be used.

1. They can be used in monitoring behavior to predict future values of forcing functions or errors, and hence aid in the detection of abnormalities.
2. They can predict future values of required outputs to compensate for disturbances.
3. They can predict the effect of actions and hence avoid feedback loop delays is evaluating the results of action.
4. They can, by means of predictive behavior, allow the system to operate open-loop and hence reduce the amount of incoming information which is processed, reducing the load on the human limited capacity single channel.
5. They can be used to predict future changes in environmental states, and hence allow programs of acton to be prepared in advance, rather than during the time that they are required, thus reducing time for adaptive changes to a minimum.
6. They can be used to optimize sampling intervals and hence maximize the number of sources of information which can be efficiently monitored.

It is clear that internal modeling is absolutely central to the efficient operation of supervisory control. During periods when the system is functioning normally the main role of the model will be to schedule sampling, and ot detect abnormalities. When an abnormal state occurs the model will aid diagnosis by predicting what variables other than the one observed to be abnormal should also be found to be abnormal, and by representing the causal structure of the process, and relations among its parts. The model will then function as an adjunct to, or substitute for a procedures manual for dealing with emergencies. (Pack, et al. 1979 notes that nuclear power plant operators have differing views as to whether more formal procedures are preferable to more flexibility and initiative. The latter requires efficient models.) Furthermore, if the operator has to come back into the control loop adequate models (transfer functions, control "subroutines", etc.) will be necessary to avoid reliance on a slow and relatively ineffective closed loop error correcting behavior involving learning anew the required contol laws.

How effectively does the human use his models for extrapolation and prediction? When the situation is simple, in the sense of there being only a single dimension onwhich control is to be excercised, he is on the whole very efficient. He can choose an appropriate control law, and even if several sources must be monitored he can build an efficient sampling schedule. Difficulty arises when he must mix modes of behavior, and particularly when conscious processes are required. It appears that conscious voluntary information processing frequently locks out model driven behavior. For example, if the operator becomes aware that a particular part of the system is in error, he will cease to monitor the rest of the environment. Such informational "tunnel vision" may occur even in automatic processing, as we saw from Cedar's (1977) results on car driving under stress, where eye movements changed so that less of the environment was sampled, and that less frequently. The reports by Weiner (1977) and Ruffle Smith (1979) on pilot errors in simulators and commercial airlines emphasize most strongly the interaction between consciously guided and model driven behavior.

Conscious prediction and exrapolation are rather inefficient. Man uses to little information and is too rigid. Kvalseth (1978) found in a Bayesian extrapolation experiment that his participants used only one sample where 4 would have been optimal, and (1978b) in a preview task that operators appeared to use only one item ahead, even when more were available. Rouse (1973) found that subjects based extrapolation only on the last three events. Their behavior could be modelled on the assumption that they applied weighting constants to these last three items, (although the weighting constants appeared to be based on much longer sequences). This suggests a mixture of unconscious procedures (evaluation of weighting constants) and conscious estimation (observation and memory of the last three items). Michaels (1979, personal communication) asked subjects to predict the next in a series of intervals and found no evidence that prediction was based on more than the last interval, even though subjects had control over the sequence generation. Gai and Curry (1978) asked subjects to observe a string of dots and to predict which side of a marker the line would pass, and found that they were strongly influenced by their initial decision, unwilling to change, and seemed to extrapolate only on the basis of the last observation. Kriefeldt (1972) suggested that in a series of responses the error was affected only by the current and the preceding responses, with no earlier information being considered. Cliff (1973) found that when tracking was disrupted by shadowing, the d.c. holds injected by the operatorwere removed at instants when the error signal crossed the position of the hold, so that no information other than instantaneous information was required to reacquire the track.

All in all one if forced to the conclusion that conscious extrapolation is inefficient. On the other hand unconscious extrapolation seems often to be more effective. Pew, Duffendack, and Fensch (1967) found that the human operator can abolish phase lag andmay indeed even introduce lead when tracking repetitive signals. He also reported (Pew, 1974) that performance improved on a section of random track which repeated every trial but which was embedded in nonrepeating segments, where performance did not improve. There are several anecdotal reports of tracking continuing for some time after a display has been turned off, and also, of course, of people who drive automobiles for considerable distances without being aware of doing so. Kelley (1968) has a lengthy discussion of prediction and of machine aiding by means of predictor displays. Although the use of the latter has not been extensive there is no doubt that they can help the human operator greatly, a fact which itself underlines the fact that he is a poor predictor. Kriefeldt and Wempe (1973) report that a predictor display improved accuracy and decreased workload in a flight simulator. Predictive displays are currently being introduced in commercial jetliners.

Theories of complex human skills such as sports have always laid great emphasis on prediction, and it is rather startling, when reviewing the literature, to come away with the feeling that when asked to predict humans are remarkably poor at doing so. On the other hand the acquisition of sports skills which require tracking, target acquisition, interception, etc., give an overwhelming impression of efficient predictive behavior. It is of course characteristic of sports skills that they are acquired by extensive practice, and that the skill becomes "automatic", that is, controlled by unconscious rather than conscious models. If this degree of efficiency is required in industrial skills, there appear to be two possibilities. One is to organize training in such a way that the operator must acquire a strongly predictive internal model, and the other is to ensure that predictive computation is available to aid the human, or that the human hands over to an automatic system. A recent study by Yoerger (1979) begins to explore such an exchange. His operators controlled simulated aircraft dynamics in several modes, which varied from completely manual control to one in which the coordinates of way-points and desired altitude were entered and navigation was automatic. He found that the effectiveness of automatic supervisory control was strongly dependent on the choice of appropriate variables. The more direct the effect of the variable which was automated the more beneficial was automation. For example, automating height or vertical speed was more effective than automating pitch angle or pitch rate. The effect of automation was to decrease the number of actions made by the operator, improve performance, increase the interval between operator intervention, and to decrease the subjective workload.

Response Generation

Although the basic mode of operation of a supervisory controller is one which puts the emphasis on monitoring rather than on controlling an automatic system, some discussion of limits on response generation is worthwhile because the supervisor must become a controller at least to fine tune a system, and may have to become a major controller when emergencies occur which the automatic system is unable to control. The vast majority of research into control has been in the context of analog control,¯ using joysticks, force sticks, roll balls, etc. But there is now need to consider limitations on other sorts ofmotor responses due to the increasingly common use of keyboards, touch panels, etc., to input digital commands.

While traditional psychological research has concentrated on isolated keyboard responses using the reaction time paradigm, more relevant work would be that on such skills as typing, (see e.g. Shaffer and Hardwick, 1968, 1969). Isolated responses are rarely if ever used in real situations. A more useful concept is that of "motor programs" (Keele, 1970; Bernstein, 1967; Pew, 1974), purposive sequences of action which can be run off as a unit when required. It is well known that controlling responses is a more demanding task than acquiring sensory information, despite the fact that the bandwidth of response mechanisms is very much lower than that of the sensory system. To that extent monitoring should be inherently easier than control.

However, the exact relation between control and supervision is by no means clear. Wickens and Kessel (1977) provide a particularly interesting discussion of the relation, and conclude that depending upon the task demands control may be easier or harder than monitoring. Among other things they point out that there are reports in the literature (McDonnell, 1966, for example) of complete adaptation to system changes without the operator being aware that there has been any change at all, presumably because kinaesthetic information was sufficient for the operator to compensate for system failure in such a way that no visual signs of error were noticed. This is almost the converse of the other well established fact that when response errors are made the human operator generally recognizes their occurrence, sometimes even before the action is complete, and corrects the error (Rabbitt, 1966; Rabbitt et al. 1978).

It is usually said that with automatisation comes a reduction in the information processing load, but it is very hard to find any explicit research on the topic. On the other hand there is good agreement that conscious response generation is a major bottleneck in the human information processor. It is of particular interest that the maximum rate at which movements can be initiated seems to be independent of the limb used. In all cases only about 1 to 3 movements per second are possible, whether the eyeball, the hand, or a leg is used for tracking. The strong presumption must be that the rate is limited by central computation in the nervous system, and not by the inertia or power/weight ratio of the limb.

There is some evidence that operators try to minimize movement as they become efficient at tasks, which would again argue for operators at least preferring supervisory control. Thus Jackson (1970) claims that all human operators minise mean square stick movement as part of their cost functional. It is well known that skilled pilots minimize the amplitude and frequency of their control movements, and the finding by McRuer and Weir (1969) that increasing the amplitude of a forcing function in the crossover region has a greater effect on work load than increasing the signal bandwidth may relate to this. Hess (1977) also found that the workload was related to the amount of control motion required, and Ashkenas (1966) found that rapid accelerations are particularly hard to control. In general it seems that violent responses make a control task difficult.

If that is the case, then supervisory control should be very much easier, especially in complex tasks with many respones and in systems with a high bandwidth. However, even in supervisory systems where moment to moment analogue control is not required the human operator is certainly "in control" in some sense. It is only that he is not required constantly to emit movements. This other sense of control is the relevant one, for example, when considering Krol's claim (1971) that ATCs find there to be less of a load when they are controlling an aircraft than when they merely observe its track on radar. This almost certainly relates to the importance of prediction in reducing workload. When the ATC has an aircraft under control his commands ensure that except for occasional aberrations he can predict its future position and state vector for many seconds after he has issued a command. When he is merely observing the flight path there is more uncertainty as to what the pilot is doing, and hence the monitoring task is harder than the "control" task.

The limitations of man as a manual controller have been well documented over the past 20 years. Both classical and optimal control theory have been used successfully in a variety of situations to predict the human operator's performance, although most of the studies have been restricted to systems with less than three degrees of freedom. Effectively the human has a bandwidth of about 1 Hz, above which he reduces his gain and shows "regression". If we consider discrete movements the so-called "psychological refractory period" limits him to less than three responses per second even in simple tasks, although in preprogrammed sequences of movement, such as typing or playing musical instruments much higher rates can be achieved, and with special keyboards such as Stenotype and Palantype word rates of several hundred a minute are possible with exceptional operators, in sustained bursts. (This incidentally suggests that there is a great deal of research still to be done on optimal data entry systems, and also on methods of training. In a keyboard which used chord inputs, for example, Sime (unpublished) found that by using adaptive training all subjects were better than the mean of the nonadaptively trained group, and the variance was much smaller. There is no doubt that current data entry devices are far from optimal.)

The well known limitations of man in controlling unstable systems or high orders of control may be related to some of the properties we have already observed. Since high order sytems induce phase lag they effectively produce a load on running memory if the operator tries to discover the effect of specific responses, and in so far as he tries consciously to control the system he will lock out his more efficient automatic responses.

Since the observed variable is continuous he will have even greater difficulty in identifying which events are due ot his actions and which are due to the forcing function when there is considerable phase lag. An unstable system requires close monitoring since any delay in responding to an error signal will mark the beginning of divergence. Subjectively both high order of control and unstable elements are associated with high subjective difficulty for the operator, (Moray, 1979a,b).

The above approaches to modelling the human operator are well understood, and call for little comment. But they are likely to be increasingly rare as automatic systems become more sophisticated and supervisory control more common. More revealing of the kind of problems caused by response generation are the scenarios described by Ruffle Smith (1979) and Weiner (1977). Their reports describe sequences of events in which tere was not obvious overloading of the operators in terms of high bandwidths, or unduly high orders of control. Very often there do not even seem to have been excessive time stresses. But extremely maladaptive sequences of behavior occurred in what were essentially supervisory situations, and of course it appears that a similar problem caused at least some of the difficulties during the Three Mile Island accident (Spectrum, 1979).

It was suggested earlier that central conscious computation and high precision were both aspects of a task which resulted in the operator being occupied for long periods, and rendered the automatic sampling schedules inoperative. The Ruffle Smith and Weiner reports emphasize this. Once the supervisors have become consciously concerned with an abnormal situation there is a pronounced tendency to "lock up" and process information only from a small part of the array. Part of this effect may be due to the fact that variables tend to be correlated, not independent. It is not clear whether this should be a help or a hindrance to supervisory control. On the one hand, the fact that a variable is observed to have some particular value should reduce the information processing required to establish the value of other variables which are correlated with it. Since the correlation establishes an a priori probability of the other variables' ranges of values, less information will be needed to estimate the actual value, and the correlation should help. On the other hand it is possible that once an error has been detected the effect of correlation is different. The knowledge of correlation will suggest that if an abnormal reading is present on the originally observed variable, then the causal structure of the system will render it likely that other abnormal readings are to be expected in a certain subset or subsets of the system variables. This will be of assistance in diagnosing the failure since it will limit the range of variables to be examined. But equally it will perhaps prevent the supervisor from examining variables which are not in the set which he regards as being the most probable variables to be associated causally with the one on which he has observed an error, and hence his sampling will be restricted to an undesirable small subset of the system variables. If then abnormal states occur in variables which are not thought by the operator to be highly correlated with the initial error they will go unobserved and uncontrolled. As mentioned before it seems that conscious control has priority over automatic control, and conscious decisions over those made by the internal model. If so, then again it suggests that automatic systems would be furnished with some way of tracking the sampling behavior of the human operator to make sure that he has not become locked up in a suboptimal pattern.

A GENERAL LOOK AT TASK INTERFERENCE

This area of psychology is covered under the general title of "attention", and many reviews exist of research over the last 20 years. Among the works which should be consulted are Broadbent, 1958; Broadbent, 1971; Moray, 1969; Kahneman, 1973; Rolfe, 1977; Sanders, 1979; Jax and Clement, 1979; Wickens, 1979; and other articles in Moray , 1979. The most widely accepted conclusion is probably that the brain is composed of a number of subsystems, each of which is single channel but some at least of which can be used in parallel with one another. Conscious behavior is strictly single channel.

What follows is a brief review, designed not to be exhaustive, but to show typical results in the area.

There is abundant evidence from laboratory studies that even under the most optimal sensory inputs may interfere with one another if they are in the same modality. For example, Moray et al. (1976) and Ostry et al. (1976) found that when listeners who were very highly practiced detected the occurrence of auditory signals which were masked by noise or near threshold the detection of a target on one channel reduced the detectability of targets on another channel. This was true for a wide variety of auditory materials. But with slightly more detectable signals listeners appeared to process signals in parallel. These experiments also found evidence that listeners were changing their response criteria as rapidly as twice a second, and that over a period of about ten hours of exposure to the messages the response criteria approached the optimal. We have already noted that because of the structure of the visual system there must be interference when different fixations are required to read different instruments, and that integrated displays may be expected to reduce the difficulty, but at the cost of introducing other kinds of interference (such as that due to different targets appearing in the same physical location and an increase in memory load if information from several sources is to be integrated). Savage, Wierwille and Cordes (1978) showed this in a more "realistic" laboratory study in which the observers had to monitor several meters and also a display of visually presented digits, finding that the efficiency with which the digits were processed depended on the number of meters. (A general treatment of the influence of the visual angle of a display with the efficiency with which the visual field is sampled will be found in Sanders, (1963) in which he reports on the relative size of the fixed visual field, the "eyefield" which can be scanned by moving the eyes but without moving the head, and the "headfield" which can be scanned by using both head and eye movements. This monograph does not appear to be as well known as it deserves from a practical point of view. For some reason many psychologists who study attention have not felt that the notion that the size of a display limits attention to be interesting. The writer does not share their view.)

Interference within central cognitive processes is commonly observed, and a study by Treisman (1965) will suffice as evidence. She computed the information transmission rate at which listeners could repeat aloud An ongoing prose passage while they heard it, and compared the transmission rate achieved by bilingual listeners who translated the message simultaneously into their second language. The information transmission rate was considerably lower when translation was involved, despite the fact that the entropy of the messages in the two languages was virtually indentical. Although her subjects were not professional translaters, it is well known that even the latter have limited ability to perform simultaneous translation without a written text to guide them. Such central transformtaions cause a marked reduction in the performance of transmisison tasks, and the effect is certainly related to the reports of Ruffle Smith and Weiner. Another interesting case is a study by Baddeley (1966) in which he found that as participants were required to perform increasingly difficult card sorting tasks, "random" number strings which they simultaneously generated became more redundant. On the other hand Moray and Harrison (unpublished) found no change in the redundancy of random number strings generated, but an increase in error on a mental arithmetic task, which suggests that the aspect of the task which the operator believes to be the more important determines the allocation of processing capacity. For a recent review of secondary task effects see papers in Moray (1979b), and Ogden et al., (1979).

The situation with regard to interference between responses is confused. On the one hand, as has been mentioned, there is wide agreement that it is on the response side that the human operator is most limited. On the other hand the most intensively studied tasks have been tracking tasks, and it is not clear under what conditions response interference is found in such tasks. Psychological studies of tracking have been relatively few, and have used rather unsophisticated methods of analysis until very recently (although see Poulton, 1974, for a different opinion). Engineering studies of tracking have used more sophisticated analyses, but have frequently been very defective from the point of view of experimental design. Senders (unpublished) reported that in a dual axis tracking task a very fine grain analysis showed that there was a negative correlation between the accuracies on the two axes over periods of a fraction of a second, and suggested that at any moment the operator was controlling either one or the other axis, but not both. Wickens, 1976 and Wickens and Gopher,. 1977, could not find any evidence that when tracking was shared with shadowing either task locked out the other. Van Lunteren (1979) opted for a parallel processing system in controlling two tracking tasks, and found slight crosstalk between them. Levison (1979) and other workers following his OCR approach (e.g. Wewerinke, 1974) argue for competition by control processes among a total capacity, implying strong interference among different responses. Broadbent (1958) pointed out long ago that doubling the rate of information transmision by doubling the rate of presentation on a single channel. Hess (1977) reported that task difficulty depended on the number of observed variables to be controlled, and Sperandio (1971, 1978) that as the number of aircraft to be controlled rose, the number of variables processed decreased.

On the other hand one can find an equal, or nearly equal number of studies where ther is little or not interference between tasks. Moray et al. (1976) found that when judgments had to be made of changes in two messages, one involving intensity and the other frequency there was much less interference tan when both messages required intensity judgments. Allport, Antonis and Reynolds (1972) found less interference between two messages one of which was speech and the other music or visual, than when both messages were speech. Underwood (1974) found that with prolonged practice a listener could detect targets in two messages as well as in one, and Moray et al. could only find interference effects near threshold. (The latter finding may be a clue to some of the discrepancies. It may only be in laboratory tasks with every short or near threshold signals that some of the limits on processing become noticeable. Where the signal-to-noise ratios are high, in "everyday" settings, there may be more than enough redundancy in the signals to conceal the interference. One might speculate that detecting tigers in presence of snakes would how interference in the noisy environment of a jungle, but little interference in a zoo.) Wempe and Baty (1968) found that when an auditory detection task was paired with a tracking task the information transmission rate on the tracking task fell, but the net transmission of the two tasks was greater, another common finding. The transmission rate was affected by the order of control, which again points to the importance of central effects, but not to the conclusion that one task precludes the other. Zeitlin and Findleman (1975) found that random deigit generation did not interfere with tracking, and McLeod (1977) similarly found that a keyboard response interfered with tracking, but a spoken response did not. Roldan (1978 unpublished) in a bimodal tracking study found no evidence for direct interference between two tracking tasks, in that there was not negative cross correlation between the error scores, although overall performance was lower with two than with one task. Apparently the fact that in several of these studies different modalities are used tends to reduce or abolish interference, as predicted by Treisman's Analyser Theory (Treisman, 1969). Cliff (1973) using a different method of analysis from Wickens came ot similar conclusions, finding that while tracking sometimes stopped during shadowing the operator appeared to monitor the forcing function, so that one cannot say that one task locks out the other. Indeed Cliff's results provide evidence of quite subtle control during time sharing. His operators tended to introduce d.c. holds in their tracking at moments when the magnitude of the error was small, and then concentrate on the shadowing task, but acquire enough information about the track to ensure that they began to track again approximately as the target crossed the value of the hold, so that movements to reacquire the target were minimized (a finding which again suggests that operators prefer to minimize movements.

At least two studies actually point to conditions where two tasks are better than one. Pitkin and Vinje (1973) found that redundant visual and auditory inputs aided a Jex critical task, and Krol (1971) in his study of air traffic control of a single aircraft found that the effect on a secondary task was less when controlling than when monitoring, a result which we have already seen is probably due to the predictability of the task under control. He also found that there was not difference in task difficulty depending on whether or not the pilot performed a secondary task, a finding which one cannot but feel is highly task dependent in view of the reports of Ruffle Smith and Weiner!

Altogether the picture is confused and confusing, and seems at best to be directly relevant only to the moments when supervisory control has given way to direct intervention in the system. There is so little actual control exercised during the supervision of a normal process that these studies on tracking cannot be very directly relevant to our understanding. Basically the competition during supervision must be between different sources of information seeking to attract the operator, or between information acquisition and internal cognitive processes used to organize and interpret that information. It would seem that with a well designed system motor responses must play a very small part. The only exception to this would be if the data entry systems (keyboards, light pens, etc.) put high demands on the operator when calling up parts of the display, and although there is little evidence on the point, it would seem likely that touch panels, providing their resolution is adequate, should have a marked advantage over other methods, due ot the inherent "compatibility" with normal pointing movements.

Time-sharing as a Skill

A recent development of interest is the reserarch on whether there is a specific time-sharing skill which can be learned. In one sense there obviously is: it is possible to learn the bandwidths of observed processes and to change one's sampling schedule. But recent work has concentrated on a different question, whether performance on a single task is a good predictor of dual task performance, or whether dual task performance involves some quite different skill. Hicks and Wierwille (1979) claim that either subjective judgments or primary task performance are better predictors of dual task performance than is performance on a secondary task, suggesting perhaps that "effort" or "concentration" is the main factor. Wickens & Tsang (1979) also found a high correlation between subjective difficulty and primary task, but not secondary task performance. But Gopher and North (1977) and North and Gopher (1976) found that primary task performance was not a good predictor of dual task performance. They noted large individual differences in the way in which operators improve on dual task performance, and that operators are very much affected by their understanding of the relative importance of the tasks, which suggests that studies in this area may be highly task specific.

Damos and Wickens (1979) and Jennings and Chiles (1977) argue for the existence of a specific time-sharing ability which can be increased by training, and also point to the papers by Moray et al. (1976) and Ostry et al. (1976) in which $\underline{d}'$ for divided attention changed without changes in single channel $\underline{d}'$ as evidence for their view. Jennings and Chiles also suggest that data driven tasks are more affected by time sharing than are cognitive tasks, which would fit with the findings about field-dependence mentioned earlier. But it runs counter to the general feeling that central and response factors are more important.

A number of points could be investigated in an attempt to clarify the picture. Firstly, and particularly inview of the changing technology in man-machine systems, the emphasis on tracking should perhaps be reduced. While it has obvious "ecological validity" in the context of conventional vehicular control and even process control, it is becoming ess important. And while there are very powerful mathematical methods of analysis both in the frequency and time domains for handling tracking data, they may well miss certain aspects of the human operator's performance. For example, tracking is a "continuous task". But does the operator track or monitor continuously? Both Senders et al.'s and Sheridan's approach would argue that he does not. And the fact that the bandwidth of the sensory system is so much higher than that of the motor system, and the bandwidth of the controlled processes on the whole rather low, further suggests that the human operator may well not be fully occupied by any single channel continuous task, since at around 1 Hz the autocorrlation function should guarantee reasonable prediction of even a random signal for a second or more. Most of the conventional measures (Bode Plots, Transinformation, etc.) are long term averages, and we are here concerned with moment to moment fluctuations. Quite marked departures from continuous performance can go undetected in such averages. The writer once simulated the d.c. holds mentioned by Cliff, and found that even when they constituted about 15% of the record they were virtually undetectable by conventional analysis, and Bekey (personal communication) simulated a sampled data system and found extreme difficulty in identifying the presence of sampling as soon as sampling interval varied by as little as 10%. If tracking is to continue to be relavent to the study of supervisory control new methods are needed for analysing the data (see for example Enstrom and Rouse, 1977).

There is no doubt about where research is needed: in the area of multiple tasks rather than dual tasks, and on a time scale and in a situation which bears more resemblance to real working conditions. Rasmussen (1979) noted that as an operator becomes skilled there is a change in the kind of model of the human operator which is needed. Initially we need a theory of how the operator processes information. But the experienced opeator is almost transparent, the properties of the process appearing through his behavior, slightly noisy and perhaps filtered by some of his motor limitations. To investigate these kinds of changes requires experiments with properties much closer to Ruffle Smith's simulation than to conventional laboratory experiments. It seems from the work of Weiner and Ruffle Smith that the "transparency" of which Rasmussen talks, and which characterizes the human operator at his best, is a property of what we have called internal models. when the operator has to rely on conscious cognitive processing the transparency is lost and there is massive interference between different aspects of the task. There seems to be a cognitive interrupt system which overrides the automatic system and, *on the timescale involved and for the kind of tasks involved in real supervisory control* makes the human operator into a severely limited capacity single channel system. While there may not be much interference between tasks in simple laboratory tracking experiments, there is massive interference in situations where the different aspects of the task have value and meaning, and where conscious decisions must be made, causality discovered, and emergency actions taken. The aim of design for supervisory control must be to provide information relevant to the handling of those sorts of situations, and to the return of the operator to an appropriate phase of his normal behavior at the end of the emergency. Procedural information needs to be provided in such a way as to minimize the time for which the operator is locked out of his "transparent", automatic behavior.

THREE SPECIFIC PROBLEMS

This review will conclude by looking briefly at data on three specific problems in complex man-machine system interaction: error detection, air traffic control, and the origin of mental workload.

Error Detection

The kind of error detection we are here concerned with is the detection of abnormal values of a continuous o digitally displayed state variable, not the tracing of faults in circuits and such like (for which see e.g. Rouse, 1979). The task of the supervisor is to detect a departure of the mean or variance of the variable from its normal values. All the evidence points firmly to the fact that the human operator makes use of the magnitude of the error and the first derivative of the error when monitoring analogue displays. Not much work seems to have been done with digital displays.

Some studies relevant to this problem have already been cited during the discussion of prediction and response generation (Wickens and Kessel, 1979; Phatak and Bekey, 1969; Moriarty, 1974). There are a number which suggest that the operator decides that an abnormal situation is present when the error and its derivative, $\underline{e}$ and $\underline{e}'$, are about 1 standard deviation away from their normal values. For example, Niemala and Krendel (1974) looked for the area in the $\underline{e}$-$\underline{e}'$ phase plane where man needed computer adssistance, and found results which can be represented by Figure 5. Onstott and Faulkner (1978, 1979) have predicted human error detection successfully on the basis of an "Urgency function",

$$U = a|e| + b\left(\frac{e \, . \, e'}{|e|}\right)$$

where a, b are constants. The function takes into account both the magnitude of e, e' and whether the error is increasing or decreasing. A large value of U is likely to mark the beginning of an abnormal situation, and therefore would be a signal for the system to override the normal sampling schedule and lock onto the particular variable showing the large U. Qualitively U would provide switching lines in the phase plane as shown in Figure 6. It may be that the tendency of experienced operators to intervene less frequently than novices when controlling a system is due to their discovery of the "no action required" regions of the phase plane. A similar approach has been suggested by Wang (1975) for quality control using a mean-variance phase plane. Repperger (1979) reports that there are characteristics loci on the e-e' phase plane which describe good and poor trackers. If the locus is small with respect to the forcing function, and the variations of the locus are normally distributed, then the tracker is good. In so far as the automatic systems in supervisory control function as the trackers, in principle such an approach could be mechanised and used to present a "probability of failure" signal directly to the supervisor. Repperger claims that his method is only applicable to deterministic forcing functions, such as sine waves or aircraft flight paths (which he considers deterministic piecewise because the movements of the aircraft are strongly constrained by the flight dynamics, and atmospheric turbulence effects are small compared with the speed of high performance aircraft). But this seems perhaps unnecessarily pessimistic, since many processes are probably at least as "deterministic" as flight paths, or can be modelled as a deterministic signals corrupted by low power noise.

It is clear that the approach to error detection through the pase plane approach is very popular. The greatest weakness seems to be that no systematic studies have been made of how operators use the phase plane under different conditions. For example, we have seen that most of the literature suggests that a departure of about one standard deviation is required for the operator to identify a failure. This cannot be an absolute value, and a systematic study is needed of the effects of probability, payoff, and instruction on the locus of switching lines in the phase plane. There is clearly an argument for developing phase plane displays for operators rather than displaying the values of the variables directly, and also for going further and using the e, e' information to display the "probability that there is an error" directly to the supervisor, giving him the option subsequently of asking either for the phase plane display or for the raw values of the variables. Given the availability of computing power there is no reason why such an error probability display should not be based on correlations between variables, and bypass some of the information processing load which is otherwise imposed on the operator by the diagnosis of complex interactions.

A few other papers may be mentioned. Wohl (1970) suggested that the phase plane display of first and second moments of a time series against each other provides a display which is very sensitive to changes in the time series, (compare Wang, 1975). Two other papers we have already quoted are by Jagacinski and Miller (1978) on the use of the root locus plot to discover how accurate a model an operator has, and Gonzales and Hovington (1977) on an algorithm for error detection which is better than human operators. Henrikson and Fu (1969) suggested computing running values of non-parametric statistics of sequential observations to detect failures, an approach essentially similar to the suggestion made above of a "probability or error" display. It would be very simple to implement the Mann-Whitney or Kolmogorov Smirnov test with a microprocessor on-line. Indeed one can envisage a system in which the control computer keeps a record of state variable values and displays statistics and predictions based on them, and also urgency functions and probability of error, while the human supervisor can set confidence limits above which the system should call his attention in accordance with his experience. If the human operator does make extensive use of sequential decision algorithms, then he is likely to be prone to "locking up" onto a critical variable, and the ability to pass to the computer the task of collecting evidence until some confidence level is exceeded and only then displaying the information should provide a powerful means of reducing his load and gaining time for higher level decisions and diagnoses.

Another area for research is in the effect on error detection of different control laws. The number of studies of failure detection is not, in fact, very large. In addition to studies such as those of Wickens and Kessel (1979) on the role of different kinds of sensory input, other parameters should be studied. For example Ince and Willeges (1974) studied the detection of a change from a control law of 1/s to 1/s . The change in dynamics could occur at a rate of 3%, 6% or 9% per second. At the higher rate a greater proportion of change occurred before it was detected, which is counter-intuitive. It is possible that the rapid build up of error was interpreted as a transient in the forcing function rather than a change in the controlled dynamics, or that at high ds/dt the correlation between successive observations was so low that the operator simply waited to try to form a clear picture of what was happening. The result suggests that an extensive program of research into the effects of hard and soft failures and their interaction with controlled element dynamics might repay the effort in establishing the natural history of error detection and putting it on a firmer foundation.

Several workers have suggested the use of queueing theory in the analysis of fault detection. For example, Walsen and Rouse (1978) required pilots to fly a simulator and check through a fault tree when a fault was discovered. Queuing theory predicted successfully the interaction of the fault diagnosis with other aspects of the task. In view of the data in Ruffle Smith and Weiner it is clear that queuing theory holds out good prospects for a description of conscious decision making, probably by assuming random arrival times, and service times related to payoff matrices, signal-to-noise ratios, etc. Fault tracing seems to be rather difficult for humans, although most of the laboratory studies have suffered from the fact that the problems have been entirely abstract, which means that the operator's knowledge about the world cannot be used. There is an interesting discussion by Pack, Seminara, Gonzalez and Parsons (1979) concerning the control of nuclear power stations. Operators apparently disagree as to whether clearer and more detailed procedural manuals are to be preferred to flexible innovation by the operator. It seems unlikely that for very large and richly interconnected systems the human operator can hope to have a detailed understanding of the system. Hence procedures are an absolute necessity. But since in most cases abnormal situations begin with a failure in one or a few variables, and and emergency only develops if the error propagates through the system, it might well be that providing the initial failure can be detected the human would be able to control a "local disturbance" without recourse to massive documented procedures. The tradeoff between the two approaches is obvious. The overall complexity means that the operator requires procedural information. But the scale of that information may be so great as to render it almost useless. (Ruffle Smith describes a flight engineer balancing twelve sheets of paper on his lap during an emergency, and shows a picture of 20 square metres of paper representing a flight plan. See also Hopf-Weichel et al. 1979.) However the information is presented it seems likely that once conscious problem solving and decision making is involved, and procedures must be consulted, there is a very real danger of th operator falling progressively further and further behind the events, and losing contact progressively with the values of the state variables as his attention is pre-empted by the conscious problems. A major design effort should be aimed at providing information, and prompting the operator, in such a way as to allow him both to work at the conscious level, and, simultaneously, to avoid the loss of system-wide status information. Any incorrect diagnosis will compound the problem rapidly, since it will lead to him inspecting increasingly inappropriate parts of the displayed information. This in turn, taken with his known unwillingness to change his hypotheses, will lead to a breakdown of control. Computer aiding should at all times be aimed at moving the operator ahead of the system in time, so that he is driving the system and not responding to it. Perhaps something akin to computer aided medical diagnosis is required, with a probabilistic estimate of the likely source of the trouble. By emphasizing the probabilistic nature both to accept it when appropriate, but also to feel free on occasion to reject the proposed diagnosis. In so far as there will always be more ways a system can go wrong than the designer has thought of, the latter option is most important.

Air Traffic Control

We are concerned here with the constraints on the control behavior of the air traffic controller, not with the system performance (how many aircraft a given airport can handle per hour, etc.). The situation is inherently one of supervisory control, but is unusual in that there are two levels of control exercised by humans. The ATC exercises supervisory control over the pilots, who in turn exercise manual or supervisory control over the flight of the aircraft. An excellent data base is available in a paper by Soede and Coeterier (1971), who provide a time line analysis with histograms of the proportion of time spent on different subtasks at a major European airport. The most commonly used model for ATC has been queuing theory, which is really the theoretical justification for time line analysis. Schmidt (1978) modelled ATC workload with queuing theory, and suggested that there was an effect analogous to "compliance", in which increasing the load produced no effect until it was quite heavy, whereupon a sudden failure would occur. Stassen and Sheridan (1979) made a similar suggestion in a general discussion of workload. Phillipp, Reiche & Kirchner (1971) suggested that the difficulty of ATC is mainly due to time stress, (equivalent implicitly to adopting a queueing theory approach), and found that difficulty appeared to depend not in the information content of the messages handled per se, but rather on the amount of time between the end of one message and the beginning of the next, which is what one would expect from a single channel system rather than a limited capacity system.

Unlike the tasks which we have examined so far, ATC seems to be a task which is characteristically under conscious control, which may account for the rather small number of aircraft which can be handled by a single controller. Sperandio (1971, 1978) estimated that seven aircraft is the absolute upper limit that a single controller can handle, which if anything is surprisingly high in view of our remarks earlier about the small span of dynamic memory. Extensive use is of course made of memory aids, such as flight strips, "shrimp boats", and in more modern displays of computer generated status information displayed on the radar. Nonetheless Sperandio found that as the number of aircraft increases controllers reduce the number of state variables per aircraft that they monitor effectively. Furthermore, behavior becomes stereotyped at high loads. This appears to be because the controller will constrain the aircraft to fly "by the book" so as to make its performance more predictable and relieve him of some of the decision making. (At very low loads behavior is also not very variable, and this appers to be the opposite side of the coin. When there is time enough the ATC will let the flight pattern be determined more by the requirements of the aircraft, and this will be similar for similar aircraft. At intermediate levels there are more ways to solve the scheduling problem, and a greater variation is seen in the solutions adopted by different controllers as they choose a compromise between their desires and those of the pilots.)

As mentioned earlier, it would seem worth investigating this tendency further, in the context of other tasks. If the tendency to stereotype at high levels of difficulty is general, then training becomes of very great importance, so as to ensure that adaptive stereotype is seen in emergencies rather than just "freezing". It seems as though a general description of the human's response to varying loads might be to say that under normal conditions of low load and when everything is proceeding smoothly the properties of the system dominate the behavior of human operators. As the load increases there comes a point where there is a maximum degree of variability, beyond which the properties of the human operator as a limited capacity single channel conscious processor dominate the behavior of the system, as he tries to keep the system variability within bounds which are tolerable for him. That is, at either low load or high load there will be little variability in the behavior of the human operator, but at low load the system causes the behavior to be sterotyped, while at high load the man is responsible. If severe overload occurs, the human will try to handle only a subset of the system variables, and progressive loss of control will occur.

ATC is a heavily "cognitive" task. Probably only the scheduling of eye movements is unconscious, and that only sometimes. (See Table 1 and Figure 2 above). Bisseret (1971) suggested that the ATC keeps track of the relationship of pairs of aircraft, and updates his estimate from time to time of their relative future positions. He remarks that it is an unusual skill in that the tolerated error rate is in fact zero. And in the light of the report by Ruffle Smith, (1979) this suggests that there must be a considerable number of errors detected and corrected before they have time to affect the system. (One cannot avoid the feeling that there is a major need for research on this point. We have estimates of performance errors, but the apparently low rates may conceal a much higher rate of behavioral errors which are made but corrected.) In the absence of electronic or paper and pencil aids there is some evidence that memory loss is quite rapid. Loftus (1979) found a 15% loss of information in 15 seconds, and Moray and Richards found that the memory for the location of a simulated radar echo was described by an equation of the form

$$SD_t = g + 0.02(t)^{1.5}$$

where g is a constant whose value depends on the experience of the controller, t is the time since an observation was made, and SD is the standard deviation of the estimate of the position of the echo. The value of g also depends on the number of echoes, and other properties of the display, and is in the range 5 - 15. The standard deviation is in mm on a 30cm diameter display. This gives significant loss beginning very near to the time given by Loftus.

Computer aiding is extensively used to reduce the load on the ATC, and more sophisticated systems are being developed. There can be little doubt that anything which reduces the memory load will be of assistance, with the proviso that good graphics are most important, in order that the time taken to read information is not increased due to poor legibility. It may be that memory for the information will be poorer with computer generated displays than with conventional displays, since there is some (weak) evidence to suggest that material is better remembered if an operator takes active steps to record it, rather than merely reading it, a point to be born in mind in general in designing displays. Hopkin (1971, 1979) has observed that when computer aiding is introduced into an ATC system the operators do not merely perform the task as before but with less load, but may radically alter the wasy in which they perform the task. As with many complex and richly connected systems, a modification to the system does not leAve one with the original system slightly modified, but with effectively a new system. A minimal design requirement would be that a new display reduces the service time for each item while keeping accuracy and workload constant.

The Origins of Subjective Load

The whole point of introducing automatic systems and moving man from the role of an active controller to that of a supervisor is to make the control of very complex systems easier and more realiable - to reduce the load on the human operator. What is the nature of the load?

Four possible meanings seem to be distinguished in the literature. Subjective Load, Imposed Load, Performance Load and Physiological Load (Stress). There is no general agreement on the terminology. Physiological stress in what is measured by such variables as heart rate, respiration rate, catecholamine excretion, etc. Performance load is measured by the occurrence of errors or the increase in latency of response. Imposed load is defined in terms of task variables, such as simultaneoity of signals, information content, signal-to-noise ratio, etc. Subjective load is what the operator experiences. From a practical point of view the most important of these measures is Performance Load, since when that becomes too great errors in system performance occur. From the operator's viewpoint the most important is subjective load. It should be noted that there is no reason to expect there to be a high correlation between them. An operator might feel quite relaxed while making a poor showing as a controller, or may feel very harrassed at a time when his performance is virtually perfect. Indeed we would expect the latter to be the case if the operator tries harder as the task becomes more difficult. Just before performance breaks down he will find the task subjectively extremely difficult but show no errors.

A more complete review of the literature on subjective load is available (Moray 1979b), and here we will merely summarize the relation between the nature of supervisory tasks and what is known of the origins of subjective load.

A human supervisor is in the following situation:

> By long experience he has built up (an) internal model(s) of the statistical structure of the system, including an estimate of causeal and correlational links among the variables. This model controls his sampling of the variables, and takes into account payoffs, meaning and value. The models also control overt behavior in well practised motor skills.
>
> The payoffs include such features as the probability of missing important events on channels other than the one being currently sampled, and the probability of faults on each variable.
>
> The sampling behavior is controlled unconsciously and tends to be optimal within the limits of the human's ability to estimate statistics, and the quality of the displays. Man is poor at estimating rare events and low bandwidths.

This unconscious sampling behavior, which is the heart of supervisory behavior, is in potential conflict with conscious behavior over which the operator has conscious control. It is also in conflict with sequential decision algorithms which determine how long a source must be sampled in order to acquire a good estimate of the valuable of a state variable displayed on it. If acquired information seems, due to the observed values of $\underline{e}$, $\underline{e}'$, to indicate a likely failure there will be conscious intervention and the sampling algorithm will be over-ridden, the operator locking onto a subset of variables which his models believe are relevant to the error.

Any conscious use of memory interferes badly with sampling. Memory only stores between one and three items of information and is constantly being over-written and updated. Extrapolation into the future is used when possible but is based on three or fewer past values.

When behavior is automatic and under unconscious model driven control responses tend to be few and of small magnitude. This is not true of voluntary responses. Even the simplest of the latter are limited to about 3 per second, and 1.5 Hz for analogue control.

Man is not good at using mixed programs. For example, if two elements have to be controlled with different transfer functions. Conscious control inhibits model driven behavior whether covert (sampling) or overt, and two unconscious models cannot be simultaneously used unless they can be transformed into a single model.

The following features of tasks are known to produce feelings of high subjective load:

1. Time pressure in the sense of concurrent demands on a time line.
2. The requirement to generate lead when controlling elements with high order transfer functions.
3. Instability in the controlled element.
4. Heterogeneous dynamics in the controlled elements.
5. Large amplitude signals in tracking.
6. Noisy signals and high precision in either perception or response.
7. Boredom.
8. Physical hard work.

All these factors have been found to be relevant to the way in which operators answer the direct question "how hard do you feel this task to be?"

In view of what we have seen about the limits on human information processing, they are not surprising. The basic fact that man is similar to a single channel system, and the probable use of sequential decision making account for (1)x and (6). (3) may also be related to this - a divergent element will require almost constant monitoring to prevent the rapid growth of error. (4) and (5) are simply empirical observations, and while (5) is somewhat puzzling, but may relate to Fitts Law which predicts that precise movements will take longer than course movements, and large movements will take longer than small ones merely by reason of mechanical inertial. Overall there seems to be considerable reason to concur with the opinion of Senders (1979, personal communication) that the major source of difficulty is time stress, and anything which can be done to reduce this will make the task of the human operator easier.

RECOMMENDATIONS WITH REGARD TO INSTRUMENTATION AND RESEARCH IN SUPERVISORY CONTROL

The recommendations which follow are made on the assumption that there are two types of behavior shown by human operators in supervisory control. The first is when the system is normal. The second following the discovery of an abnormal state of at least one variable, whether the latter is detected by the computer or the human operator.

When the system is normal a well practised operator will behave automatically, and use a nearly optimal scanning strategy, of the details of which he is unaware. The major problem here is the potential conflict between the sampling (scheduling) algorithm and the data acquisition algorithm. Display design must be directed to minimizing dwell time (data acquisition time). In this mode it is to be expected that the probability of successive samples being taken from highly correlated data sources will be low, and sampling will therefore adequately survey most of the sources.

When an abnormal state is found the behavior of the operator will be under conscious control, and will be subject to human limitations of short term memory, extrapolation, conscious decision making, and inadequacies as an "intuitive statistician". In this mode it is probable that successive samples will be taken from highly correlated sources, since most systems will be compound of tightly coupled subsystems, and an error in a part of a subsystem means that diagnostic information should be sought in the first instance within the subsystem. This will have the effect of producing "cognitive tunnel vision", with the result that large parts of the array may go unexamined for long periods. It is possible that the only effective way to cope with this problem will be to have one operator deal with that part of the system which has been discovered to be abnormal and another to deal with the (perhaps temporarily) still normal part of the system. In this way the nearly optimal ability of the operator in scanning the normal system can be used without interuption while conscious decision processing is used to handle the abnormality. The recommendations about displays made below are in the light of the need to help the conscious processing of information and prevent cognitive tunnel vision.

1. Displays must minimize data acquiaition time. Providing that adequate ergonomics are used, VDU's should be efficient, but size of display, s/n ratio, color coding, contrast, and glare, etc. must be considered. The fact that the display area is small is good.

2. Displays must minimize memory load and the need for significant extrapolation. For these purposes VDU's which display only parts of the system are bad. The system must prompt the operator if he ignores some subsystems and does not call them up for long periods.

This implies that the system must keep a log. It may be advantageous to have the system adapt itself to the strategy used by the operator and then test to see whether it is optimal in a queueing theory sense, and interact with the operator to help him improve.

3. It is worth investigating the idea that what displays should show is the probability of error, rather than the value of the state variables. If so it seems that the best way to display them will be on an (e,e') phase plane with switching lines indicated at probabilities associated with an urgency function. There is every reason to have the system do these computations and force a display when the switching lines are crossed even if the operator does not call for it. Predictor displays should be implemented where possible. On a phase plane display it may be worth leaving a history as well, like a radar "tadpole". With color displays the past, present and future could be displayed in different colors. Other possibilities include displaying the probability that a variable is normal, or "star diagrams" or "snow flakes" to display multivariate error probabilities.

4. Control actions should be tolerant of errors. There is now increasing evidence that in man-machine-systems ther are frequent errors of behavior, but that the systems usually filter them out, because the response time of most complex systems, either due to physical inertia (aircraft) or the time constants of physical processes (process control) allow the operator to recover and correct himself. There is good evidence that in the case of many if not most errors the operator knows that he has made a behavioral error even before he completes the action. This does not mean that all actions should be queried. It may be that having the system model the error history of the operator would allow the former to query the latter on actions with frequent corrections.

5. A major effort will have to be put into the ergonomics of the new displays, from the point of view of keyboard layout, functions available, optical properties, etc. (For example, on the DEC VAX keyboard, the system can appear to go dead because the "no scroll" key is where a relaxed operator is likely to rest the ball of his left hand when resting between entries. One does not want the system to appear to go dead during emergencies.)

6. The problem of presenting procedural information to be used during fault diagnosis, start up, etc. is a major one. This includes the optimal allocation between initiative and algorithmic sequences. How does one accomodate 20m of procedures (Ruffle-Smith, 1979)

efficiently within the conscious working space, (neural, electronic, or paper), of the user? This is of vital importance, since as we have seen, the conscious processes of data handling and computation pre-empt all other information processing, including the computing of "urgency functions".

7. There is need for computer aiding in fault diagnosis once an error has been detected. It is important to prevent the operator from becoming locked onto one hypothesis, and to aid in the representation of "causally connected subsets" of the system wherein a solution to the diagnosis is most likely to lie. Probably the system should keep a record of its past faults and perform some kind of pattern matching, displaying the results probabilistically, but preventing the operator from accepting a very probable solution without checking it.

8. Fast time modelling of the system should be available. This has two advantages, firstly that the operator may explore system properties. Secondly, during fault diagnosis he may make a stochastic estimate of the likely consequences of his chosen interventions. This is especially important where the time constants of the system are such that a transient injected may take a long time to affect the process but even longer to die out. Humans are bad at handling long lags.

9. As a general rule, the aim should be to train the operator to levels of efficiency where the vast majority of his normal interaction is done "automatically", and without conscious symbolic thought. When he himself becomes the process he is efficient. When he has to think about it he is not. On the other hand enough conscious interaction must be retained to prevent boredom and promote a feeling of responsibility.

It may be advantageous in emergencies to adopt a team approach. The fault tracer would work consciously on the subsystem showing abnormalities, while another operator worked "automatically" on those parts of the system which still appeared normal.

With the last recommendation we move into the realm of individual values and role playing, which is outside the scope of the present paper, but important without a doubt if man-machine systems recommendations are to be about the man and not merely about the machine. It is as well to remember that at all times and for all systems the fault you have not forseen will sooner or later occur; and the fault for which the man has not been trained on is the one he will have to control.

References

Allport, A.; Antonis, b.; Reynolds, P. 1972. On the division of attention: A Disproof of the Single-channel Hypothesis. Quart. J. Exp. Psychol. 24.

Ashkenas, I.L., 1966. Some open and closed loop aspects of lateral directional handling characteristics. AGARD 533.

Attreave, F. 1953 Applications of Information Theory & Psychology. Holt. N.Y.

Audley, R.J. 1973. Some observations on theories of choice reaction time. In Kornblum, S. (ed.) Attention and Performance IV. Academic Press, N.Y.

Baddeley, A. 1966. The capacity for generating information by randomisation. Quart. J. Exp. Psychol. 18. 119-129.

Bainbridge, L. 1978. Forgotten alternatives in skill and workload. ergonomics, 21, 169-185.

Baker, C. 1963. Further towards a theory of vigilance. Vigilance: A Symposium. ed. Buckner, D. & McGrath, J. McGraw-Hill N.Y. 127-153.

Baron, S. & Berliner, G. 1977. The effects of deviate internal representation in the optimal model of the human.

Baron, S., Kleinman, D. & Levison, W. (1970). An optimal control model of human response. Automatica, 6, 357-369.

Baxa, E. & Nolts, L. 1972. Adaptive signal detection with finite memory. IEEE Trans SMC-2, 42-49.

Bernstein, N.A. 1967. The co-ordination and regulation of movements. Pergamon. London.

Bisseret, A. 1971. Analysis of mental processes involved in air traffic control. Ergonomics, 14 565-570.

Broadbent, D. 1958. Perception and communication. Pergamon. London.

Broadbent, D. 1971. Decision and stress. Academic Press. N.Y.

Carbonell, J. 1966. A Queueing Model for many-instrument visual sampling. IEEE Trans. HFE-7, 157-164.

Cedar, A. 1977. Driver's eye movements as related to attention in simulated traffic flow conditions. Human Factors, 19, 571-583.

Chiles, D. & Alluisi, E. 1979. On the specification of operator or occupational workload with performance measurement methods. Human Factors, 21, 515-528.

Clement. W. F., Jex, H. R., & Graham, D. 1968. A manual control-display theory applied to instrument landings of a jet transport. IEEE MMS-9, 93-110.

Clement, W., Hofmann, L., & Graham, D. 1973. A direct procedure for partitioning scanning workload in control display system design. 9th Annual Conf. of Manual Control. 389-400.

Cliff, R. C. 1973. Attention sharing in the performance of a dynamic dual task. IEEE Trans SMC-3, 241-248.

Coeterier, J.F. 1971. Individual strategies in ATC freedom and choice. Ergonomics, 14, 579-584.

Conant, R., & Ashby, R. 1970. Every good regulator of a system must be a model of that system. International Journal of Systems Science, 1, 89-97.

Crossman, E.R.F.W. 1959. A theory of the acquisition of speed skill. Ergonomics, 2, 153-166.

Curry, R., & Gai, E. 1977. Detection of random process failures by human monitors. in Sheridan, T. & Johannsen, G. (eds) Monitoring Behaviour and Supervisory Control. Plenum. N.Y.

Dale, H. 1968. Weighing evidence: an attempt to asses the efficience of the human operator. Ergonomics, 11, 215-230.

Damos, D. L., & Wickens, C.D. 1979. The identification and transfer of timesharing skills. (in press)

Davis, R., Moray, N. & Treisman, A. 1961. Imitative responses and therate of gain of information. Quart. J. exp. Psychol. 13, 78-90.

Enstrom, K.D., & Rouse, W.B. 1977. Real time determination of how a human has allocated his attention between control and monitoring tasks.

Fitts, P. 954. The information capacity of the human motor system in controlling the accuracy of movement. J. Exp. Psychol. 47, 381-391.

Gai, E., & Curry, R. 1978. Perseveration effects in detection tasks with correlated decision intervals. IEEE SMC-8, 93-101.

Garner, W. 1962. Uncertainty and structure as psychological constructs. N.Y. Wiley

Gonzales, R. C. and Hovington, L.C. 1977. Machine recognition of abnormal behavior in nuclear reactors. IEEE SMC -7, 717-728.

Goodenough, D. 1976. A review of individual differences in field dependence as a factor in auto safety. Human Factors, 18, 53-62.

Gopher, D. & North, R. 1977. Manipulating the conditions of training in time-shared performance. Human Factors 19, 583-593.

Grant, J.S. 1971. Concepts of fatigue and vigilance in relation to railway operation. Ergonomics 14 111-118.

Green, D.H. & Luce, R.D. 1973. Speed-accuracy trade-off in auditory detection. in Attention and Performance IV. ed. Kornblum, S. Academic Press. N.Y. 547-570.

Henrikson, E. & Fu, K. 1969. Calamity detection using non-parametric statistics. IEEE SEC-5, 150-155.

Hess, R. 1977. Prediction of pilot opinion ratings using an optimal pilot model. Human Factors, 19 459-475.

Hess, R.A. & Teichgraber, W.M. 1974. Error quantization effects in compensatory tracking tasks. IEEE SMC-4, 343-349.

Hicks, T., & Wierwille, W. 1979. Comparison of 5 mental workload assesment procedures in a moving-base driving simulator. Human Factors, 21, 129-144.

Hopf-Weichel, R., Lucaccini, L., Saleh, J., & Freedy, A. 1979. Aircraft emergency decisions: cognitive and situational variables. Perceptronics PATR-1065-79-7.

Hopkin, V.C. 1971. Conflicting criteria in evaluating air traffic control systems . Ergonomics, 14, 557-564.

Hopkin, V.C. 1979. Mental workload measurement in Air Traffic Control. in Mental Workload, Theory and Measurement. ed. Moray, N. Plenum Press, N.Y. 381-386.

Ince, F., & Williges, R. 1974. Detecting slow changes in system dynamics. Human Factors, 16 278-285.

Jagacinski, R.J. & Mille, R.A. 1978. Describing the human operator's internal model of a dynamic system. Human Factors 20, 425-433.

Jackson, G. 1970. On the optimality of the human controller. 6th Ann. Conf. on ManualControl. 723-736.

Jahns, D. W. 1973. A concept of operator workload in manual vehicle operators. Forshungsintiut fur Anthropotechnik Reprt #14.

Jennings, A., & Chiles, W.D. 1977. An investigation of time-sharing ability as a factor in complex performance. Human Factors 19, 535-548.

Jex, H. & Clement, W. 1979. Defining and measuring perceptual-motor workload in manual control tasks. in Mental Workload: Theory and Measurement, 1979. Plenum. N.Y.

Kahneman, D. 1973. Attention and Effort. Prentice Hall. N.J.

Keele, S. 1970. Movement control in skilled motor performance. Psychological Bulletin, 70, 387-403

Kelley, C. 1968. Manual and automatic control. Wiley, N.Y.

Kickert, W. Bertrand, J., & Praagman, J. 1978. Some comments on cybernetics and control. IEEE SMC-8, 805-809.

Kleinman, D. L., Baron, S., & Levison, W.H. 1970. An optimal control model of human response. Part 1. Theory and Validation. Automatica, 6, 357-369.

Kornblum, S. 1975. An invariance in choice reaction time with varying numbers of alternatives and constant probability. in Attention and Performance V, ed. Rabbitt, P. & Dornic, S. Academic Press. 366-381.

Krendel, E., & McRuer, D. 1960. A servomechanism approach to skill development. J. Franklin Institute, 269, 24-42.

Kriefeldt, J.G. 1972. A dynamic model of behavior in a discrete open loop selfpaced motor skill. IEEE SMC-2, 262-273.

Kriefeldt, J., and Wempe, T. 1973. Pilot performance during a simulated standard instrument procedure turn with and without a predictor display. 9th Annual Conference on Manual Control, 147-162.

Krol, J.P. 1971. Variation in ATC-workload as a function of variation in cockpit workload. Ergonomics, 14, 585-590.

Kvalseth, T. 1978. Effect of preview on digital pursuit control performance. Human Factors, 20, 371-377.

Kvalseth, T. 1978. Human and Bayesian information processing during probabilistic inference. IEEE SMC-8 224-229.

Kvalseth, T., 1978. Quantitative models of motor responses subject to longitudinal, lateral and preview constraint. Human Factors, 20, 35-39.

Laming, D. 1973. Mathematical Psychology. Academic Press. N.Y.

Leonard, J. A. 1959. Tactual choice reaction. Quart. J. Exp. Psychol. 11, 76-83.

Levine, J. and Samet, M. 1973. Information seeking with multiple sources of conflicting and unreliable information.

Human Factors, 15 407-419.

Levine, J., Samet, M., & Brahlek, R. 1975. Information seeking with limitations on available information and resources. Human Factors, 17, 502-513.

Levison, W. 1979. A model for mental workload in tasks requiring continuous information processing. in Moray, N. (ed) Mental workload: theory and measurement. Plenum. N.Y.

Loftus, G. 1979. Short-term memory factors in ground controller/pilot communication. Human Factors, 21 169-181.

Mackworth, J. 1959. Paced memorizing in a continuous task. Journal of Experimental Psychology, 58, 206-211.

Monk, T. 1977. Sequential expectancy in visual search. Human Factors 19, 601-609.

Monty, R.A. 1973. Keeping track of sequential events: implications for the design of displays. Ergonomics, 16 443-454.

Moray, N. 1963. Attention: selective processes in vision and hearing. Hutchinson. London.

Moray, N. 1979a. The subjective measurement of mental workload. MMS Laboratory Report, M.I.T.

Moray, N. 1979b. Mental workload: theory and measurement. Plenum, N.Y.

Moray, N. 1979c. in Strategies of information processing. ed. Underwood Academic Press.

Moray, N., Fitter, M., Ostry, D., Favreau, D., & Nagy, V. 1976. Attention to pure tones. Quarterly Journal of Experimental Psychology, 28, 271-284.

Moray, N., Synnock, G., & Richards, S. 1973. Tracking a static display. IEEE SMC-3, 518-521.

Moriarty, T. 1974. Manual control of a time varying vehicle. 10th Annual Conference on Manual Control, 535-550.

Mowbray, G. H. & Rhoades, M.V. 1959. On the reduction of choice reaction times with practice. Quart. J. Exp. Psychol. 11, 16-23.

Mullis, C.T. & Roberts, R.A. 1968. Memory limitations and multi-stage decision processes. IEEE SMC-4, 307-316.

Murthy, P. 1975. Optimal control of unreliable systems with discrete time inspection. IEEE SMC-5 452-456.

Niemala, R. & Krendel, E. 1974. Detection of a change in plant dynamics in a man-machine sytem. 10th Annual Conference on Manual Control. NASA 97-112.

North, R., & Gopher, D. 1976. Measures of attention as predictors of flight performance. Human Factors 18, 1-14.

Ogden, G.D. Levine, J.M., & Eisner, E.J. 1979. Measurement of workload by secondary tasks. Human Factors 21, 529-548.

Ostry, D., Moray, N., & Marks, D. 1976. Attention, practice and semantic targets. Journal of Experimental Psychology, Human Performance, 2 326-336.

Onstatt, E.D. & Faulkner, W.H. 1978. Predictoin, evaluation, and specification of closed loop and multiaxis flying qualities. AFFDL-TR-78-3.

Onstatt, E.D. & Faulkner, W.H. 1979. Definition and interpretation of flying qualities. TR-NOR-79-22. Northrop Aviation Hawthorne, California.

Pachella, R. 1974. The interpretation of reaction time in information processing research. in Kantowitz, B., (ed) Human Information Processing: tutorials in performance and cognition. Lawrence Erlbaum. N.J.

Pack, R., Seminara, J., Eckert, S., Seidenstein, S., Gonzales, W., Stempson, R., & Parson, S. 1979. Human factors methods for nuclear control room design. EPRI NP-1118-SY.

Peterson, C., & Beach, L. 1964. Man as an intuitive statistician. Psychological Bulletin, 68, 29-46.

Pew, R. 1974. Human perceptual-motor performance. in Kantowitz, B. (ed) Human information processing: tutorials in performance and cognition. Lawrence Earlbaum. N.J.

Pew, R., Duffendack, J.C., & Fensch, L.K. 1967. Sinewave tracking revisited. IEEE HFE-8, 130-134.

Phatak, A., & Bekey, G. 1969. Decision processes in the adaptive behavior of human controllers. IEEE SSC-5, 339-351.

Phillipp, V., Reiche, D., & Kirchner, J., 1971. The use of subjective rating. Ergonomics, 14, 611-616.

Pitkin, E.J. & Vinje, E.W. 1973. Comparison of human operator critical tracking task performance with visual and aural displays. IEEE SMC-3 184-187.

Poulton, C. 1974. Tracking Skill and Manual Control. Academic Press. N.Y.

Posner, M., Klein, R.M. 1973. On the function of consciousness. in Attention and Performance IV. 21-35 ad. Kornblum. S.

Academic Press. N.Y.

Rabbitt, P. 1966. Errors and error correction in choice response tasks. J. Exp. Psychol. 71, 264-272.

Rabbitt, P., Cumming, F., & Vyash, S. 1978. Some errors of perceptual analysis in visual search can be detected and corrected. Quant. J. Exp. Psychol. 30, 319-333.

Ragan, T.J., Back, K.T., Stansell, V., Ausburn, L.J., Ausburn, F.G., Butler, P.A., Huckabay, K., 1979. Cognitive styles, a review of the literature. AFHRL-TR-78-90(1).

Rasmussen, J. 1979. Reflection on the concept of operator in Moray, N. (ed) Mental Workload: theory and measurement. Plenum. N.Y.

Repperger, D.W., Ward, S.L., Hartzell, E.J., Glass, B.C. and Summers, W.C. 1979. An algorithm to ascertain critical regions of human tracking ability. IEEE SMC-9, 183-196.

Robinson, J.A. & Bennick, C.C. 1978. Field articulation and working memory. Journal of Research in Personality. 12, 439-449.

Rolfe, J. 1977. The measurement of human response in man-vehicle control situations. in Sheridan, T. & Johannsen, G. (eds) Monitoring Behavior and Supervisory Control. Plenum. N.Y.

Rouse, W.B. 1972. Cognitive sources of suboptimal human behavior. NGL 22-009-002 Mechanical Engineering Laboratory, M.I.T.

Rouse, W.B. 1973. A model of the human in a cognitive prediction task. IEEE SMC-3, 473-478.

Rouse, W.B. 1978. A model of human decision making in a fault diagnosis task. IEEE SMC-8 357-361.

Rouse, W.B. 1979. A model of human decision making in fault diagnosis tasks that include feedback and redundancy. IEEE SMC-9, 237-241.

Ruffell Smith, P. 1979. A simulator study of the interaction of pilot workload with errors, vigilance and decisions. NASA TM-78482. Ames Research Center.

Sanders, A. 1963. The selective process in the functional visual field. Inst. for Perception RVO-TNO. Sosterberg, Netherlands.

Sanders, A. 1979. Some remarks on mental load. in Moray, N. (ed) Mental workload: theory and measurement. Plenum. N.Y.

Savage, R., Wierwille, W. & Cordes, R. 1978. Evaluating the sensitivity of various measures of operator workload using random digits as a secondary task. Human Factors, 20 649-654.

Schmidt, D.K. 1978. A queueing analysis of air traffic controller's workload. IEEE SMC-8, 492-498.

Schrenk, L.P. 1969. Aiding the decision maker: a decision process model. Ergonomics, 12, 543-557.

Senders, J., Elkid, J., Grignetti, M., & Smallwood, R. 1966. An investigation of the visual sampling behavior of human observers. NASA CR-434.

Shaffer, H., & Hardwick, J. 1968. Typing performance as a function of text. Quarterly Journal of Experimental Psychology 20, 360-369.

Shaffer, H., & Hardwick, J. 1969. Errors and error detection in typing. Quarterly Journal of Experimental Psychology, 21, 209-213.

Sheridan, T. 1970. On how often the supervisor should sample. IEEE SSC-6, 140-145.

Sheridan, T., & Ferrell, R. 1974. Man-machine systems. M.I.T. Press.

Sheridan, T. & Johannsen, G. 1977. Monitoring behavior and supervisory control. Plenum. N.Y.

Sheridan, T. & Rouse, W.B. 1971. Supervisory sampling and control: sources of suboptimality in a prediction task. 7th Annual Conference on Manual Control.

Shiffrin, R.,& Schneider, W. 1977. Controlled and automatic processing: II. Perceptual learning, automatic attending, and a general theory. Psychological Review, 84, 127-190.

Soede, M., & Coeterier, J., 1971. Time analysis of tasks of approach controllers in ATC. Ergonomics, 14, 591-601.

Spectrum, November 1979. Special issue on Three Mile Island. IEEE

Sperandio, J.C. 1971. Variation of operator's strategies and regualting effects on workload. Ergonomics, 14, 511-577.

Sperandio, J.C. 1978. The regulation of working methods as a function of workload among Air Traffic Controllers. Ergonomics, 21, 195-200.

Stassen, H., & Sheridan, T. 1979. Definitions, models and measures of human workload. in Moray, (ed) Mental workload: theory and measurement. Plenum. N.Y.

Stassen, H., & Veldhuyaen, W. 1977. The internal model - what does it mean in human control? in Sheridan, T. & Johannsen, G., (eds) Monitoring behavior and supervisory control. Plenum. N.Y.

Tamura, H., & Yoshida, T. 1974. Model of control and the limits of control capability for human subjects. IEEE SMC-4, 482-488.

Taylor, F. J. 1975. Finite fading memory filtering. IEEE SMC-5, 134-137.

Taylor, D. 1976. Accidents, risks, and models of explanation. Human Factors, 18, 371-380.

Treisman, A. 1965. The effects of redundancy and familiarity on translating and repeating back a foreign language. Brit. J. Psychol. 56, 369-79.

Treisman, A. 1969. Strategies and models of attention. Psychol. Rev. 76, 282-299.

Tulga, K. 1978. Dynamic decision making in multitask supervisory control: Comparison of an optimal algorithm to human behavior. Man-Machine Systems Laboratory Report. M.I.T.

Underwood, G. 1974. Moray vs. the rest: the effects of extended shadowing practice. Quarterly Journal of Experimental Psychology, 26, 368-374.

Underwood, G. 1979. Strategies of information processing. Academic Press, N.Y.

Van Lunteren, A. 1979. Identification of the human operator describing function models with one or two inputs in closed loop systems. WTHD. #113. Technical University Delft, Holland.

Vickers, D. 1970. Evidence for an accumulator model of psychophysical discriminations. Ergonomics, 13, 37-58.

Wald, 1947. Sequential analysis. Wiley, N.Y.

Walden, R.S. & Rouse, W.B. 1978. A queueing model of pilot decision making in a multitask flight management situtaiton. IEEE SMC-8, 867-874.

Wang, P. 1975. On pattern recognition and quality control. IEEE SMC-5, 470-472.

Weiner, E. 1977. Controlled flight into terrain accidents: system induced errors. Human Factors, 19, 171-182.

Wempe, T. & Baty, D. 1968. Human information processing rates during certain multiaxis tracking tasks with a concurrent auditory task. IEEE MMX-9 129-137.

Wewerinke, P.H. 1974. Human operator work load for various control conditions. 10 Annual Conference on Manual Control. 167-192.

Wickens, C. 1976. The effects of divided attention of information processing in manual tracking. Journal of Experimental Psychology: Human Performance, 2, 1-17.

Wickens, C. 1979. Measures of workload, stress, and secondary tasks. in Moray, N. (ed) Mental Workload: theory and measurement. Plenum. N.Y.

Wickens, C., & Gopher, D. 1977. Control Theory measures of tracking as indices of attention allocation strategies. Human Factors, 19, 349-365.

Wickens, C. & Kessel, C. 1979. The effects of participatory mode and task workload on the detection of dynamic system failures. IEEE SMC-9, 24-34.

Wickens, C. & Tsang, P. 1979. Attention allocation in dynamic environments. University of Illinois. TR EPL-79-3/AFOSR-79-3.

Wierwille, W. 1979. Physiological measures of aircrew mental workload. Human Factors, 21 575-594.

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

Wohl, J.G. 1970. The statistical phase plane display: a new method for monitoring time series. 6th Annual conference on Manual Control. 147-164.

Yoerger, D. 1979. Man-machine performance for a simulated aircraft with a multilevel automatic control system. Man-machine Laboratory Report. M.I.T.

Young, L. 1969. On adaptive manual control. IEEE MMS-10, 292-351.

Zeitlin, L.R. & Finkleman, J.M. 1975. Research note: subsidiary task techniques of digit generation and digit recall as indirect measures of operator loading. Human Factors 17, 218-220.

END

DATE

FILMED

1-81

DTIC