Winter 1987

The Implications of Varying Levels of Task Automation on Workload

Micheline Y. Eyraud
Old Dominion University

The Implications of Varying Levels of Task Automation

on Workload

by

Micheline Y. Eyraud

A Dissertation Submitted to the Faculty of
Old Dominion University in Partial Fulfillment of the
Requirements for the Degree of

DOCTOR OF PHILOSOPHY

INDUSTRIAL/ORGANIZATIONAL PSYCHOLOGY

OLD DOMINION UNIVERSITY
December 1987

Approved by:

Glynn D. Coates, Chair

ABSTRACT

THE IMPLICATIONS OF VARYING LEVELS OF TASK AUTOMATION ON WORKLOAD

Micheline Y. Eyraud
Old Dominion University, 1987
Director: Glynn D. Coates

The present study investigated the effects of the addition of automation on task workload. Utilizing a modified secondary task paradigm, the workload which was imposed by three different levels of automation, selected from the continuum of automation on each of two primary tasks, was assessed by comparing performance on a secondary task which remained unaided in all conditions. The levels of automation under investigation in the present study were manual, intermediate, and total aiding. The primary tasks selected for investigation were a sensory-decision making task and compensatory tracking task. A long-term memory task was chosen as the secondary task. It was hypothesized that as the amount of aiding increased on the primary task such that the amount of cognitive processing required by the

individual also increased, there would be a corresponding decrease in the performance of the secondary task. This decrease was hypothesized since in the aided conditions the individual was responsible not only for verifying that the system was in performing the task satisfactorily, but was also ultimately responsible in all the conditions for performing the task manually if deemed necessary. Sixty subjects were randomly assigned to one of six conditions by factorially combining the three levels of aiding and two levels of task combinations. Each subject received three experimental sessions.

The hypothesis that the addition of cognitive workload in conditions where increasing amounts of aiding was introduced into a task situation was not borne out by the results of this study. Results of this study suggested that a significant reduction in workload can be obtained by totally automating a particular task but that there is no significant reduction in the amount of workload when aiding, which gives advice to the individual, is compared to the situation where the individual must perform the task without assistance. Additional human factors research needs with regards to the introduction of automation were also identified.

ACKNOWLEDGEMENT

The helpful comments received from members of my committee are gratefully acknowledged. I would like to note my special appreciation to my Committee Chair, Dr. Glynn D. Coates, for his helpful advice, his infinite patience, and his mastery of the art of teaching which has made learning statistics a pleasant experience.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

General Background

The system development process consists of a series of logical, sequential and overlapping phases which have been formalized by the Department of Defense (1977a, 1977b, 1977c). These phases are broken down into a series of milestones at which the particular system under development is critically reviewed before permission is given to continue further development. Since the human operator plays a nontrivial role in the functioning of these systems, at each of these milestones there are a number of activities related to human factors engineering (HFE) which must be taken into account (Price, 1980a, 1980b). Variables such as the design of technical manuals, reliability, maintainability, training, the design of controls and displays, and selection (cf., Hutchinson, 1981; McCormick & Sanders, 1982; Meister, 1976, 1982; Salvendy, 1987; VanCott & Kinkade, 1972) are concerns which must be addressed during this process in order to result in an effective system. In addressing these issues, human factors specialists have at their disposal a large body of human performance research (e.g., Bilodeau & Bilodeau, 1969; Holding, 1981), which specifies the conditions under which performance is optimized, which may be applied to these activities. Variables such as practice (cf., Crossman, 1959; Newell & Rosenbloom, 1981; Schneider & Shriffrin, 1977; Shriffrin & Schneider, 1977), feedback (cf., Moray, 1981) and the conditions under which the individual must perform these tasks (e.g., Alluisi, 1969; Jerison & Pickett, 1963; Warm, 1984) have all been noted to affect human performance and therefore must be taken into consideration when designing the human-machine interface. In addition, the amount of workload which is imposed on the human operator by specific system configurations must also be addressed early in the development cycle.

Workload Assessment

Workload assessment is problematic in that the nature of workload is complex and multidimensional. This complexity is reflected in the complex theoretical foundations of workload assessment. Moray (1979) refers to random walk theory, accumulator theory, discrete and continuous information theory, supervisor theory, queuing theory, the Theory of Signal Detection, linear control theory, optimal control theory, and adaptive control theory in his discussion of workload. In addition, physiologically oriented theories such as those of Selye (1956) might also be brought to bear on the topic. Therefore, given the numerous workload theories proposed, many definitions (e.g., Advisory Group for Aerospace Research and Development, 1978; Johannsen, 1979; Rolfe & Lindsey, 1973; Sheridan & Stassen, 1979) as well as a number of different metrics for the measurement of workload (cf., Hartman & McKenzie, 1979; Roscoe, 1978; Wierwille & Williges, 1978; Williges & Wierwille, 1979; Wierwille, 1979) exist.

One idea which has appeared in a number of definitions of workload is that there is a relationship between the workload which is imposed by the task and fatigue (e.g., Gartner & Murphy, 1976; Welford, 1953). In evaluating these definitions, psychophysiological measures (Cumming & Corkindale, 1967; Wierwille, 1979), based on the concept of activation or arousal, have been used by researchers to examine variations in physical and mental effort.

Another notion that has inspired a large body of research in workload assessment is that workload is related to the particular task at hand, and, more specifically, to the amount of attention which the individual must devote to the task. For example, Jahns (1973) maintained that "workload is the extent to which an operator is occupied by a task". Further, Cooper and Harper (1969) proposed that workload "is the measure of additional pilot effort and attention required to maintain a given level of performance in the face of a less favorable or deficient characteristic". Similarly, Brown, Stone and Pearce (1975), Clement, McRuer and Klein (1971), Eggemeier, (1981), Singleton (1971), and White (1971) emphasized the level of individual effort required to satisfy a specified set of demands. By adopting this type of approach when measuring workload, several behavioral measures (e.g., Chiles, 1982; Chiles & Alluisi, 1979; Chiles, Alluisi & Adams, 1968; Kelley & Wargo, 1967; Williges & Wierwille, 1979) exist from which inferences regarding workload can be made. Note, however, that a measure of workload is not necessarily contingent on how much an individual actually is occupied by a task, but may also reflect an individual's perception of how much they are occupied by the task (Moray, 1982). In line with this idea, methodologies such as rating scales (cf., Borg, 1978a; 1978b) and other subjective scaling techniques have been used in workload assessment. One example of such techniques is the Subjective Workload Assessment Technique (or SWAT) discussed by Reid, Shingledecker & Eggemeier (1981).

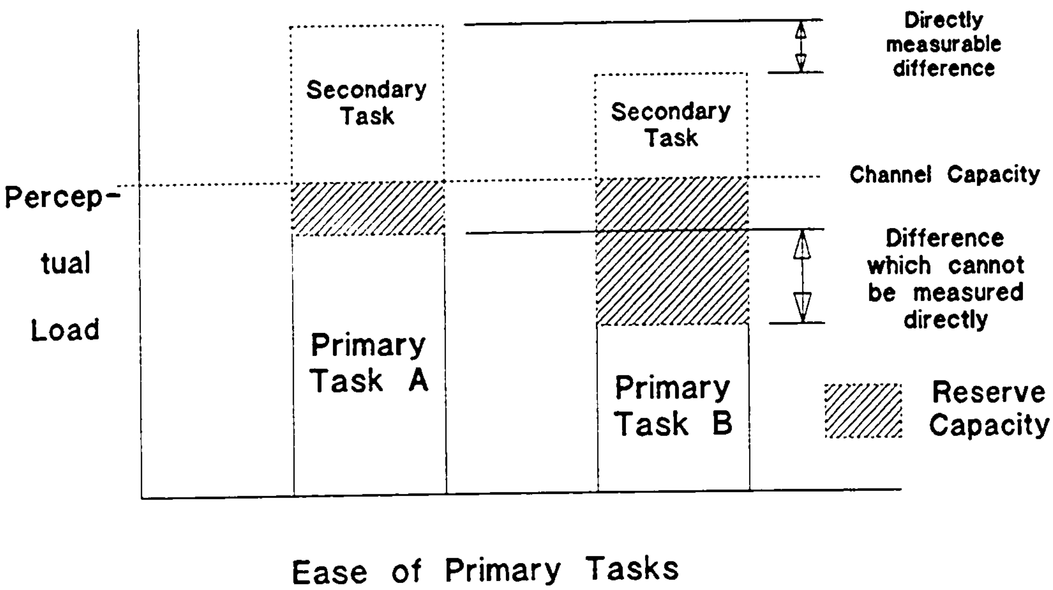
Secondary task paradigm in the measurement of workload. One behavioral methodology emphasizing attention as a measure of workload is the secondary task paradigm (cf., Brown, 1964; Knowles, 1963; Odgen, Levine & Eisner, 1979; Pew, 1979; Rolfe, 1971; Wickens, 1979). This experimental paradigm is characterized by the situation in which two discrete and separate tasks are performed concurrently with a clear emphasis on the performance of one of the tasks (the primary task). In one variation, the performance levels of the secondary task alone are compared to performance of that task in combination with a primary task. Alternatively, the performance levels of the secondary task might also be compared when variations of the primary task are introduced. In either case, performance in a secondary task is taken as a measure of the operator's spare mental capacity or the additional workload on the individual. Due to the central role played by the secondary task, there are specific requirements which must be fulfilled in the selection of secondary task: (1) the task should require little learning (to avoid practice effects), (2) it should be self-paced; (3) it should not interfere with or disrupt primary task performance, and, (4) the index of operator load that is calculated from the scores of a given loading task should be comparable from situation to situation (Knowles, 1963).

Secondary task methodology derives directly from the single channel hypothesis of attentional capacity (cf., Broadbent, 1957; 1958; 1971; Duncan, 1980; Hochberg, 1970; Kahneman, 1973; Lane, 1982; Neisser, 1970; Norman, 1968; Triesman, 1969). Capacity theories of attention, which originated from the study of the psychological refractory period (Craik, 1948; Davis, 1956; 1957; 1959; 1962; Hick, 1948; Welford, 1952), conceptualize attention as a processing resource that can be allocated as required by task demands up to a certain point, which is the individual's total capacity to perform mental work. Although the nature of this capacity has been characterized in different ways -- e.g., as the ability to transmit information

(Broadbent, 1958; 1971); to translate stimuli into responses (Welford, 1968); to assign modality specific analyzers (Triesman & Davies, 1973) or processing capacity (Moray, 1967) to the examination of stimulus input; or to expend psychological effort (Garvey & Henson, 1958; Kahneman, 1973). However, the basic assumption behind all of these theories is that tasks place a demand on a central limited capacity system which can be allocated with considerable freedom across a wide variety of tasks. Thus, faced with simultaneous inputs and limited resources, the individual must allocate relative amounts of "space" or processing capacity to the analysis of one input or another on a priority basis which results in some inputs receiving deeper, or more complete processing than others.

The secondary task paradigm provides an estimate of the difference between the "mental capacity" consumed by the main task, and the total available capacity. Thus, as the demands of the primary task increase, rendering fewer resources available to process the demands of the secondary task, performance on the secondary task will theoretically deteriorate to a degree proportional to the demand increase of the manipulated task (cf., the performance resource function of Norman & Bobrow, 1975). An illustration of the function of the secondary task is shown in Figure 1.

When examining workload from an attentional or capacity standpoint, Posner and Boies (1971) proposed that the selection of responses resulted from a competition for



This diagram, from Brown (1964), illustrates the function of the secondary task. According to the single-channel model, differences in the perceptual load imposed by different primary tasks cannot be directly measured since, by definition, the individual should not make errors until reserve capacity is exceeded. However, through the introduction of a secondary task, differences in the amount of workload imposed by the primary task may be inferred by comparing performance on the secondary task, which may be directly measured.

Figure 1. Function of the Secondary Task

non-specific resources or capacity. To assess this theory Posner and Boies conducted a study in which the primary task for subjects was to compare two letters presented in succession, one second apart, on a screen. Immediately following the second letter presentation, the subject was to press one of two keys to indicate whether the two letters were the same or different. The secondary task required the subject to press a third key in response to a brief auditory signal. The auditory signal was presented randomly during the course of a letter-matching trial. Posner and Boies found that when the auditory probe signal coincided with, or just preceded, the arrival of the second letter (i.e., overlapped with the period in which the subject responding during the primary task) there was a significant increase in the reaction time to the probe. At all other times, including the period when the subject was presumably encoding the first letter preparatory to comparing it with the second one, the response to the probe was as rapid as when the probe was presented outside the course of a letter match trial altogether.

An alternative theoretical conception of human information processing which attempts to account for attentional phenomena are the structural theories (e.g., Keele, 1973). Structural theories maintain that attention is related to the competition of tasks for specific processing mechanisms (or structures) which are necessary for performance (cf., a multidimensional resource space such as that postulated by Navon & Gopher, 1979). These theories differ from capacity theories in that they treat this competition for resources as a discrete, all-or-none process: tasks either compete for the common mechanisms or they do not, and processes either demand attention or they do not in a multidimensional performance-resource space. Thus, a structural model implies that interference between tasks is specific, and depends on the degree to which the tasks call for the same mechanisms. This differs from capacity models which maintain that interference is nonspecific, occurring when the demands of two activities exceed available capacity.

The Present Experiment

Performance in complex environments is determined in large part by the amount of information that can be effectively processed by the individual. Information processing problems such as data overload, getting lost, the inability to find, integrate, or interpret the "right" data at the "right" time are examples of the operator performance failures while using complex systems (Woods, 1986). The use of "artificial intelligence" or AI (Barr & Feigenbaum, 1981, 1982a, 1982b; Davis & Lenat, 1982; Winston, 1982) to augment operator cognitive activities seems to offer the potential to alleviate these classical performance failures.

Expert Systems Technology

One application of AI, "intelligent" computer programs, or expert systems (other terms which are used synonymously are decision augmentation systems and decision support systems or DSS) function by codifying the processes which are used in reasoning, planning or decision making in these complex situations (Gevarter, 1983a). A DSS consists of: (1) a knowledge base (or knowledge source of domain facts and heuristics associated with the problem); (2) the inference engine (or control structure) which is a set of strategies for utilizing the knowledge base in the solution of a problem; and, (3) the user interface.

The knowledge base. The characteristic of DSS's which differentiate them from traditional computer programs is that rather than relying on non-knowledge-guided search techniques or computational knowledge, the basis of an "intelligent" program is a knowledge base of the domain being aided. Feigenbaum (1982) stated:

> An "expert system" is an intelligent computer program that uses knowledge and inference procedures to solve problems that are difficult enough to require significant human expertise for their solution. The knowledge necessary to perform at such a level, plus the inference procedures used, can be thought of as a model of the expertise of the best practitioners of the field.
>
> The knowledge of an expert system consists of

> facts and heuristics. The "facts" constitute a body of information that is widely shared, publicly available, and generally agreed upon by experts in a field. The "heuristics" are mostly private, little discussed rules of good judgment (rules of plausible reasoning, rules of good guessing) that characterize expert-level decision making in the field. The performance level of an expert system is primarily a function of the size and quality of the knowledge base that it possesses. (p. 1)

In addition to declarative knowledge, DSSs also contain procedural knowledge which consists of information concerning specific courses of action. Depending upon the form of knowledge representation chosen, the two types of knowledge may be separate or integrated. Different knowledge representation schema which have been used in expert system development are: (1) state-space representation which represents the structure of a problem in terms of the alternatives available at each possible state of the problem, (2) the classical approach (e.g., Manna, 1973) developed by philosophers and mathematicians wherein knowledge is derived from the rules of formal logic; (3) a procedural approach (Hewitt, 1975; Winograd, 1972) in which knowledge of the world is contained in procedures or small programs which know how to do specific things in a well-specified situation; (4) semantic nets which were invented as a psychological model of human associative memory (Anderson & Bower, 1973; Norman & Rumelhart, 1975) and consist of nodes which represents objects, concepts and events, and links between the nodes which represent the interrelationship of knowledge; (5) direct or analogical representations (Funt, 1976; Sloman, 1975) in which there is a correspondence between the relations in the representational data structure and the relationships in the represented situation; and, (6) production systems (Davis & King, 1977; Waterman & Hayes-Roth, 1978; Winston, 1977) which integrates both procedural and declarative knowledge in the form of condition-action pairs, called productions. These productions may be either frame-based, which represent an object as a group of attributes, or rule-based systems which represent knowledge in the form of heuristic "if-then" production rules. (A more complete discussion of knowledge representation schemas is beyond the scope of the present endeavor. For further information on this topic, see Barr & Feigenbaum, 1981).

Inference engine or control structure. However, the fact that an expert system is knowledge-based does not make an expert system intelligent. In order to accomplish this, the system must incorporate another component which directs knowledge implementation. The structure of this component, termed the inference engine (also known as the control structure or rule interpreter) of a DSS is of critical concern since it is the task of the inference engine to operate on the knowledge base in such a manner that the knowledge is accessed in an efficient and consistent manner. In effect, the inference engine "runs" the expert system by determining which rules are invoked and how they are to be applied to the problem, executing the rules, and determining when an acceptable solution has been found. (A more complete discussion of control structures is also beyond the scope of the present endeavor. For further information on this topic, see Hayes-Roth, Waterman & Lenat, 1983).

<u>The user-system interface</u>. The final component of an expert system is the system user interface (Berry & Broadbent, 1987; Kidd & Cooper, 1985) which includes any part of the system which the user comes in contact with physically, perceptually or conceptually. Figure 2 illustrates the elements of this interface which are important from a human factors standpoint: the user, the task, the hardware, the software, documentation and job design issues. It is important to note that the elements of this interface and the work which will be discussed in the section which immediately follows are specifically applicable not only to DSSs, but also to human-computer interaction (HCI) in general. These two research areas, which have historically been distinct, have recently converged due to advances in technology (Gaines & Shaw, 1986).

<u>The user</u>. The user of the system is a logical starting point for any discussion of human-computer systems development. Not unlike more classical work in HF, early work in HCI focused on the classification of users so as to take advantage of certain characteristics in order to optimize the design of these systems. For example, Ramsay and Atwood (1979) suggested that computer users be classified in terms of certain abilities (e.g., sensory

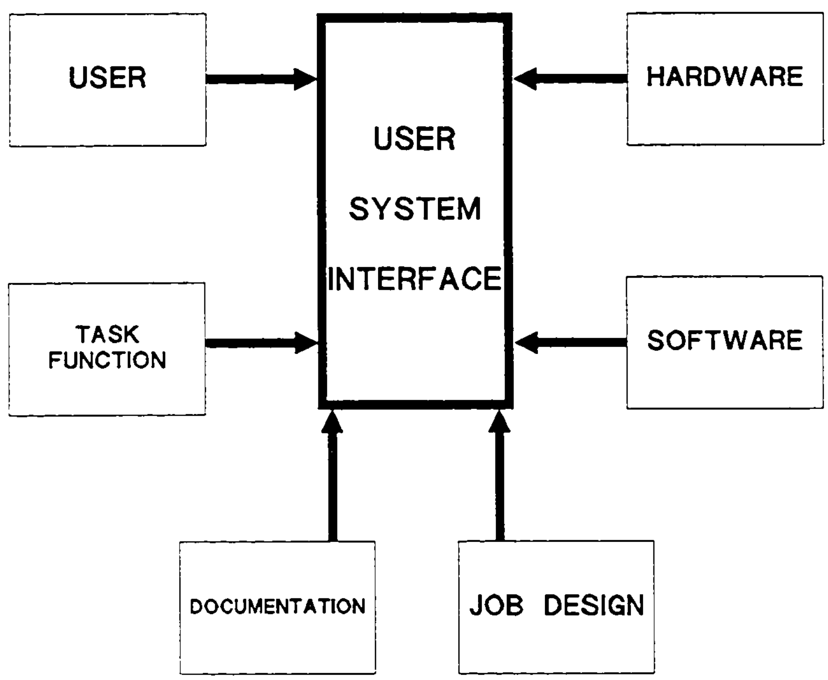


Figure 2. Elements of the User-Computer Interface

(adapted from Chapanis, 1982)

capacities, anthropometric dimensions, intellectual capacities and cognitive decision style). Another way of categorizing users focuses more on the nature of the user's job in order to produce categories such as: analysts, operators, and service personnel (Smith, 1981); clerical workers (Stewart, 1974); managers (Eason, 1974); nonprogrammers and programmers (Martin, 1973), and technical users (Ramsay & Atwood, 1979). Other ways users have been classified into distinct groups are on the basis of their familiarity or sophistication with computers (e.g., Al-Awar, Chapanis & Ford, 1981; Barnard, Hammond, Morton & Long, 1981; Bennett, 1979; Shneiderman, 1980) and in terms of the role of the human in the human-computer system - which, according to Williges (1982), might take the form of student, operator/analyst or programmer.

The emphasis on incorporating user considerations into DSSs have taken a different approach. Rather than arbitrarily characterizing the user along static dimensions, an approach which may be faulty in that it fails to account for the possibility that the individual using the system will undergo a transition over the course of time (Rich, 1983), work specifically applicable to DSSs has emphasized the development of "mental models" (Gentner & Stevens, 1983; Moran, 1981). Mental models are defined as cognitive representations of the system's internal mechanics (Halasz & Moran, 1983). These representations might be: (1) in terms of the structure of the device; (2) an inferential representation which, given the structure of the device, determines its function; (3) in terms of a description of the functioning of the device; and, (4) a representation in terms of the behavioral functioning of the system (DeKleer & Brown, 1983).

The incorporation of mental models in DSSs enable interaction between the user and the system to be tailored to the preferences, abilities, capabilities, etc. of the individual users at any particular point in time during system operation. This dynamic interaction might be accomplished in one of two ways: (1) the user might be given the capability to modify the system to suit themselves; or, (2) the DSS might be provided with information on how the individual uses the system (e.g., Self, 1977) or stereotypes (Rich, 1979) that so the system can take charge of its own personalization.

Software. The impact of computer technology on the individual is dependent to a large extent upon the demands imposed upon the operator by the type of software with which the individual must interact. In this regard, evaluations of the HCI with command languages, programming languages, dialogue systems (Williges & Williges, 1982; 1984) and error management (cf., Allen, 1984) have been evaluated.

When discussing the impact of implementing computer aiding in various situations, it is important to define the nature of the automated system under consideration.

Various individuals (e.g., Alter, 1980; Hart & Sheridan, 1984; Rouse & Rouse; 1983; Zachary, Wherry & Glenn, 1981; Zachary, 1986) have all detailed taxonomies of automation. Although the terminology varies across authors, a common theme is that there are various gradations of "expert assistance", ranging from systems which enable the operator to select from among well-defined alternatives by predicting the consequences of each alternative and computing its subjective utility at one end of the continuum to systems which exhibit initiative of their own at the other end of the continuum. An example of this type of automation continuum is shown in Figure 3.

The foundation for subdividing the DSS domain into these classes is based upon one distinguishing characteristic -- the type of decision function performed. The first difference among the different classifications lies in the level of "intellectual" processing of the underlying algorithm. A second difference among the various types lies in the nature and degree of the operator's involvement in the decision process. Near one extreme the human uses data provided by the DSS as an input to a decision, but is still responsible for the final decision. However, at the other extreme, the system is structured to make decisions on it's own, leaving the human responsible for monitoring and managing the actions of the DSS by acceptance or rejection.

Task/Function. Another element of the user

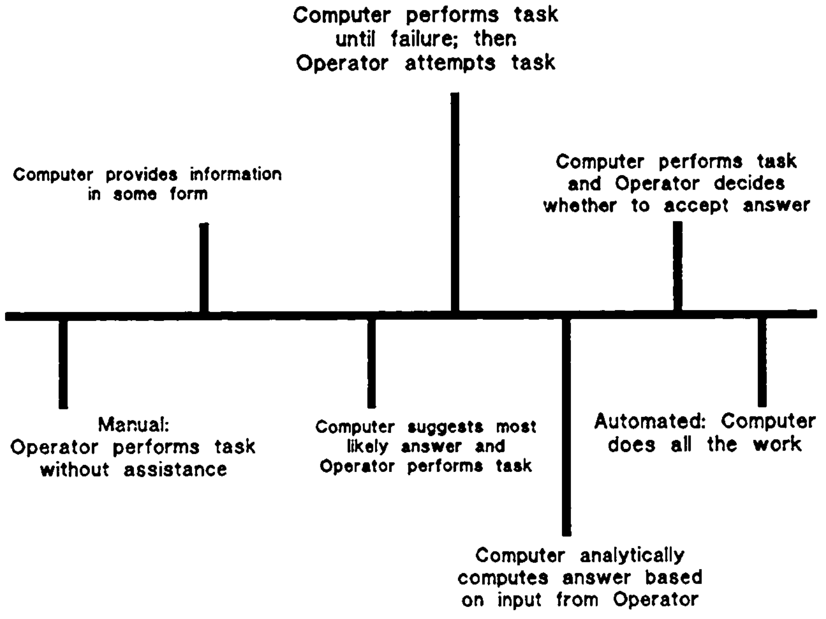


Figure 3. Degrees of Automation

(taken from Gordon, 1985)

interface is the task or function which must be accomplished by the individual/system (cf., Carter, 1986). The complexity of the task, the kinds of information needed to perform the task, and the constraints under which the task must be performed are all relevant considerations in the human factors design of computer systems. A synthesis of the types of tasks to which DSSs have been applied is contained in Table 1.

It is important to note, however, that in DSS design is the fact that the nature of the task which the human operator is required to perform changes as a function of the particular expert system. Hollnagel and Woods (1983) pointed out that with increasing automation the functions performed by the individual are changing in emphasis to include increased responsibilities for more cognitive activities with a corresponding decrease in the responsibility for the manual control aspects of the task. Sheridan (1976, 1984) noted that there are five types of tasks which the individual is required to perform in the situation where automation has been implemented. These tasks are: (1) planning, which is carried out in anticipation of a response to failure events rather than to immediate control requirements; (2) teaching, which consists of the specification of commands to the computer to make it run automatically; (3) monitoring, which is carried out on a continuous basis to insure that everything is working properly; (4) intervening, in which the operator is required to interrupt the programs

Table 1. Functional Categories for Expert System Applications

Category	Problem Addressed	Types of Systems
Interpretation	Infers situation description from data	Image and acoustic analysis, surveillance
Prediction	Infers likely consequences of given situations	Weather forecasting, crop estimation
Diagnosis	Infers system malfunctions from observations of potentially noisy data	Medical, electronic
Design	Configures objects which satisfies particular requirements	Circuit layout, chemical synthesis
Planning	Creates a program of actions to achieve goals	Automatic programming, chemical synthesis
Consultation	Provides advice regarding optimal course of action	project management
Monitoring	Continuously interprets signals and sets off alarms when intervention is required	Nuclear power power plant regulation, patient respiration
Debugging	Prescribes remedies for malfunctions	Computer software
Repair	Executes a plan to administer a prescribed remedy	Automobile, computer
Instruction	Diagnoses, debugs and corrects student behavior	Tutorial, remedial
Control	Interprets, predicts, and monitors system behaviors	Air traffic control, battle management

(adapted from Mischkoff, 1985)

and directly interface with the task or process which was previously aided; and, (5) learning from experience with the system in order to perform the other functions more efficiently.

Hardware. To date, much of the emphasis in HCI concerns the development of hardware specifications upon which to base design decisions. For example, there is an abundance of information concerning visual display hardware (e.g., Tannas, 1985); input/output devices such as keyboards (Noyes, 1983; Ramsey & Atwood, 1979); and, other peripherals such as joysticks and mice (Card, English & Burr, 1978). With specific regard to DSS's, new technologies such as computer vision (Gevarter, 1982; Marr, 1982) and natural language understanding (Gevarter, 1983b; Lehnert & Ringle, 1982) will impact the way the human operator interacts with the system.

Documentation. How the individual learns about the computer system is another aspect of the user interface which must be taken into consideration. Prior to the extensive use of computers this problem was handled through the use of printed matter external to the system which described or explained how that system worked or should be used (cf., Johassen, 1982). However, with advances in computer capabilities, the form that these instructions take and the way in which these documents are handled and stored have changed dramatically (Cohill & Williges, 1985). With DSS development, new forms of documentation designed to be interpreted by computer programs and restructured is essential in order to optimize individual performance.

Job Design. Work with regard to job design in HCI research has centered around workstation design issues (cf., Grandjean & Vigliani, 1980). However, when considering the incorporation of automation, job design issues such as the possibility that automation might impose additional levels of monotony or stress (Thackeray, 1980) as well as limit the intrinsically motivating aspects of work (Parsons, 1985; Sharit, 1985) must also be taken into account.

HFE Considerations in the Development of Expert Systems

Although there is evidence that individuals responsible for the operation of complex systems favor the increased use of automation to enhance performance (e.g., Kibbe & DeVere, 1987), it is important to remember that the incorporation of DSS's also pose a number of questions not adequately addressed by human performance research applicable to traditional human-machine systems (cf., Boehm-Davis, Curry, Weiner & Harrison, 1981; National Research Council [NRC], 1983; 1984). For example, since it may not be assumed that software which is constantly changing is completely reliable (e.g., Zakay, 1982) the effects that the lack of practice has on the ability of the individual to assess the situation effectively and to intervene when necessary is of critical concern. Price (1985) identified allocation of function as a major issue which needs to be addressed in order to optimize performance using automated systems. Although the most expedient solution for function allocation might be to assign those tasks which cannot be automated to the monitor of that system (Rouse, 1981), this alternative is problematic in that it may result in a degradation of performance since the individual would have only an incoherent set of bits and pieces of tasks to perform (Bainbridge, 1982). Perhaps the most acceptable solution to this problem would be to develop systems in which there is a dynamic distribution of tasking (Rouse, 1977; Greenstein & Revesman, 1986; Vaughan & Mavor, 1972). Other issues of importance in the design of human-computer interfaces is the development of human reliability and error analyses for the supervisory control situation; the optimum design of human-computer dialogue design (e.g., Williges & Ehrich, 1984); and, the optimum degree of "transparency" or ease with which the operator can obtain information concerning any decisions made by the system (Woods, 1985).

One common assumption is that with the automation of functions which were once relegated to human control, the workload of the human operator will be reduced and that processing resources will be freed to deal more effectively with other aspects of system requirements (Swedish, 1983). However, given that there may be additional cognitive demands inherent in the supervisory control situation (e.g., Sheridan & Johannsen, 1976), one important issue which must be considered in the design of automated systems is whether particular levels of automation which are incorporated into that system impose additional workload on the operator and whether this workload is significantly different from the task loading present in the manual control situation. Since different levels of automation function to remove quantitatively different amounts of manual task responsibility yet at the same time impose increased cognitive demands on the individual, it can be argued that the additional mental operations required with increasing amounts of task automation might act to increase the demand for the limited processing resources of the operator.

The purpose of the present study was to assess workload levels imposed through the introduction of varying levels of computer aiding on different tasks. Using the secondary task paradigm, three different levels of aiding were introduced on two different primary tasks. The different aiding levels which were studied were: (1) the manual mode in which the operator was responsible for accomplishing the task without assistance; (2) an intermediate aiding condition in which the individual was given advice on how to complete the task but made the final decision regarding the applicability of the advice before completing the task; and, (3) total automation in which the operator was responsible for verifying that the system made the correct inputs to complete the required response. Performance on a secondary task, which in all conditions remained unaided, was compared in order to assess the amount of workload which was imposed by the introduction of aiding. It was hypothesized that as the amount of aiding was increased on the primary task such that the amount of cognitive processing required by the individual also increased, there would be a corresponding decrease in their performance of the secondary task. This decrease in performance was hypothesized since in the aided conditions the individual was responsible not only for verifying that the system was performing the task satisfactorily, but was also ultimately responsible in all the conditions for performing the task manually if it were deemed necessary. Further, it was hypothesized that this decrease in performance would vary as a function of the particular task under consideration.

METHOD

The paradigm utilized in the present study represents an amalgamation of the traditional secondary task loading paradigm (cf., Chiles, 1982; Knowles, 1963; Ogden, et al., 1979; Rolfe, 1971) and synthetic work methodology (Alluisi, 1969; Chiles, et al., 1968; Morgan & Alluisi, 1972). It should be noted that there are limitations to the secondary task paradigm. For example, the paradigm cannot account for instances in which manipulations of the objective characteristics of a primary task will produce changes in the performance of one secondary task but not another (North, 1977), nor can it account for the fact that simple changes in processing structure may dramatically alter the degree of task interference of the information processing demands of the single task components (Kantowitz & Knight, 1976). However, not enough is known concerning the implications of task automation to proceed directly with alternative workload assessment techniques which take into account the theoretical construct of multiple resources (Wickens, 1980; Wickens, Montford & Schreiner, 1981).

Design

The design of the present study employed a factorial combination of three levels of aiding, two levels of task combinations and three sessions with subjects nested in the aiding and task combination factors but factorial to the session factor. There were six groups of subjects of ten each, in which each subject received only one combination of task and aiding levels for three sessions. The two task combinations involved a single secondary task paired with one of the two primary tasks to which the aiding was applied.

Independent Variables

Performance Tasks. In the present study three tasks were used -- a sensory decision making task (SDMT), compensatory tracking task (CTT) and a long term memory task (LTMT). These tasks were developed from similar, although substantially more complex tasks of the Multiple Task Performance Battery (MTPB) which was designed to tap the behavioral functions required of operators of complex systems (Alluisi, 1967). It is also important to note that the primary tasks (i.e., SDMT, CTT) were representative of different aspects of information processing. Functionally, the SDMT focuses on the integration of information and the subsequent choice (cf., Jennings & Chiles, 1977), while the primary focus of the CTT is on the resultant motor action. These tasks were selected for use in the present study since they required the use of skills which are representative of the skills or functions which might be replaced by automated systems (Defense Advanced Research Projects Agency, 1983).

Primary Tasks. At one level of the task combination factor, the SDMT was the primary task. During each experimental session, subjects were presented with one of three symbols (a "*", "@", or "#") on each of 400 trials. To each of these symbols the subjects were instructed to make one of three responses: (1) the presentation of an "*" required the subject to enter "Z"; (2) the proper response to a "#" sign was "C"; and, (3) the proper response to the presentation of a "@" symbol was "X".

For the other level of task combination the primary task was a CTT. In performing the CTT, subjects were required to monitor the movements of a cursor and make control inputs using the "<" and ">" keys to maintain the position of the cursor in the center of the tracking bar.

Secondary Task. In both levels of task combination the LTMT (Bahrick, Noble & Fitts, 1954; Hellyer, 1962; Peterson & Peterson, 1959) was the secondary task. The task for the subject was to keep a mental running total during the experimental session. At random points throughout the session the note "ENTER CURRENT VALUE" was presented on the right-hand side of the CRT screen. Subjects were required to subtract mentally a value of three from the running total and input the revised total into the computer. At the start of the first experimental session the running total was set at 1000, and for subsequent sessions the starting point for the LTMT decreased by 100 to reduce the possibility that the subject could rehearse their responses before the start of subsequent sessions.

Levels of aiding. From the continuum of expert aiding the following three levels of aiding were chosen for analysis: (1) manual (i.e., no aiding), (2) intermediate aiding in which the system makes recommendations regarding the optimal course of action; and, (3) total aiding where the system performs the task for the subject. The implications for task performance at each decision aiding level are summarized in Table 2.

Subjects and Experimental Groups

Sixty male and female students at Old Dominion University served as subjects for the study and were randomly assigned to one of six groups. The experimental groups were formed by combining primary task with level of automation (i.e., manual, intermediate or total). Subjects received course experimental credit in return for their participation.

Apparatus

Hardware. An H-89 microprocessor with CRT display was used to present all tasks and to record responses. The three tasks were displayed on distinct portions of the CRT. The LTMT and SDMT were presented side by side in the lower

Table 2. Implications of Automation on Subject Task Requirements

Sensory-Decision Making Task (SDMT)

Manual condition - Subjects are required to perform the decision-making task without assist.

Intermediate Aiding - Subjects are required to verify the advice which is presented by the computer before making their response.

Total Aiding - Subjects are required to monitor the presentation of symbols and verifying that the responses made by the computer are correct. In instances where the subjects determine that the response made by the computer is correct, they are tasked with pressing the space bar. If, however, they determine that the response is incorrect, the subjects are tasked with making the proper response as well as pressing the space bar.

Tracking task (CTT)

Manual condition - Subjects are required to make control inputs to keep a cursor within a specified area on the screen without assist.

Intermediate Aiding - Subjects are required to make control inputs to keep a cursor within a specified area on the screen given feedback as to the potential effects of their control inputs and status information on the position of the cursor.

Total Aiding - Subjects are required to monitor the performance of the computer in accomplishing the tracking task. Should the computer allow the cursor to go out of tolerance subjects must intervene and make the necessary control inputs. In addition, at random points during the experimental session, subjects are also required to press the space bar when prompted to insure their continued attendance.

two-thirds of the screen while the CTT was displayed in the upper third of the screen. A typewriter keyboard attached to the microprocessor was used to enter responses. The keys designed for the required responses were clearly designated. All remaining keys were deactivated (i.e., pressing them did not have any effect).

Software. Task presentation and summary feedback after each trial were controlled through BASIC software which was designed to simulate the functioning of expert systems which offer varying levels of expert assistance. The specific functions performed by the "expert system" for each of the tasks at each of the levels of task automation are detailed in Table 3.

Procedure

Prior to the collection of data a proposal for research was reviewed and approved by the Department of Psychology's Human Subjects Review Committee.

Each of sixty subjects participated in three one-hour sessions over a period of three consecutive days. During each of the three sessions the primary task and the level of aiding remained constant. An additional variable, that of multiple sessions, was introduced in order to insure that performance measures gathered on each of the tasks was a reflection of an individual's ability to perform the task rather than their learning of the task. Prior to the

Table 3. Implications of Automation on System Functioning

<u>Sensory-Decision Making Task</u> (SDMT)

Manual condition - System presents symbols for subject decision and response.

Intermediate Aiding - System presents symbols for subject decision as well as advice on how to respond to the presentation.

Total Aiding - System presents symbols and makes a response which may or may not be correct.

<u>Tracking task</u> (CTT)

Manual condition - System presents tracking task to the subject.

Intermediate Aiding - System presents tracking task to the subject. In addition, a numerical indication of the present position of the cursor as well as information concerning the future position of the cursor is presented to the subject.

Total Aiding - System presents tracking task to the subject and keeps cursor within a specified range throughout experimental session.

<u>Long Term Memory Task</u> (LTMT)

Manual - System presents a prompt to the subject to input the next value.

experimental sessions, each subject was briefed as to the nature of the experimental tasks and procedures and was asked to sign an informed consent form. See Appendix A for a listing of the instructions given to each subject. Following completion of the experimental sessions, each subject was thoroughly debriefed and all questions were answered.

RESULTS

Measures

As discussed earlier, each of the experimental sessions involved performing two tasks -- a primary task (either SDMT or CTT) which differed across conditions as a function of the amount of automation introduced, and a secondary task (LTMT) which remained unchanged across conditions. Performance measures gathered for the SDMT, which were calculated on-line, were reaction time (RT), standard deviation (SD) of RT, and error rate. In addition, summary data for each trial was also recorded which consisted of the trial number, time into the experimental session, the coordinates of the screen where the symbol was presented, the correct response as well as the response which was input by the subject. Performance measures for the CTT, which were also calculated on-line, were mean absolute error, RMS error, and the minimum and maximum cursor position value throughout the session. Performance measures gathered for the secondary task (LTMT) were mean RT, SD and the number (N) of trials which were responded and the number of trials which were missed. In addition, summary data for each trial was also recorded which consisted of the time at which each of the trials was presented, the correct response, the response given by the subject, and the response latency.

Primary Task Performance

For the secondary task paradigm to be a viable workload assessment technique it is imperative that performance on the primary task remain unchanged across conditions. Therefore, before mentioning results of the secondary task analyses, it must be established that subjects did not perform significantly differently across the aiding conditions.

Performance on the primary task was assessed by comparing the number of errors and mean RT across the SDMT conditions and, across the CTT conditions, RMS error, the most reliable estimate of the dispersion of a sample of discrete measures (Kelley, 1969). The results of this analysis for the SDMT, summarized in Table 4, indicate that while there was a significant effect for session (df = 2,54; p <.05; F = 36.94 and 17.04), no significant differences in performance as a function of the aiding conditions were noted (df = 2,27; p <.05; F = 1.33 and 2.53). Similar results, summarized in Table 5, were obtained for the CTT performance. While there was a significant effect for session (df = 2,54; p <.05; F = 13.47), no significant differences in performance as a function of aiding were noted (df = 2,27; p <.05; F = 3.19).

Secondary Task Performance

The secondary task in each of the six conditions was a

Table 4. Sensory Decision Making Task (SDMT) Performance

Performance Measure: Error Rate

Source	Sum of Squares	df	Mean Square	F	p
Session	4125.9556	2	2062.9778	36.94	.0001
Aiding	310.5556	2	155.2778	1.33	.2809
Sess x Aid	108.7778	4	27.1945	.49	.7453
Subj(Aid)	3146.0667	27	116.5210		
Sess x Subj(Aid)	3015.9333	54	59.8827		
Total	10707.9333	89			

Performance Measure: RT

Source	Sum of Squares	df	Mean Square	F	p
Session	1.1813	2	.5907	17.04	.0001
Aiding	.9430	2	.4715	2.53	.0984
Sess x Aid	.1946	4	.0487	1.40	.2452
Subj(Aid)	5.0323	27	1.0065		
Sess x Subj(Aid)	1.8717	54	.0347		
Total	9.2230	89			

Table 5. Compensatory Tracking Task (CTT) Performance

Performance Measure: RMS Error

Source	Sum of Squares	df	Mean Square	F	p
Session	572.0214	2	286.0107	13.47	.0001
Aiding	392.8706	2	196.4353	3.19	.0571
Sess x Aid	348.1902	4	87.0476	4.10	.0057
Subj(Aid)	1662.6499	27	61.5796		
Sess x Subj(Aid)	1136.3970	54	21.0444		
Total	4122.6992	89			

mental arithmetic task which required the subjects to count down by three from a specific number at random points during the experimental session. In addition to the measures mentioned above, preliminary analysis of the data indicated a number of distinctive errors made by subjects across experimental conditions. Additional performance measures used in subsequent analyses are described below:

Repetitions (REP) occurred when the subject made the same response as on the previous trial.

Missed (MISS) responses occurred when the subject failed to make a response on a particular trial.

Incomplete (INC) responses occurred when the subject failed to input all three numbers.

Significant Departures (SIG) occurred when subjects deviated by more than 25 from the previous response, or where subjects completely lost their place in the number sequence and were required to start over.

Tests of homogeneity of variance using the Cochran - C procedure were conducted for each dependent measure for each of the three aiding conditions. The results of this analysis indicate that for both of the task combinations the hypothesis that the variances in the two groups are equal must be rejected for all dependent measures except SD. Alternative methods of transforming the data (i.e., no transform, a square root and log transform) were compared. Results of this analysis determined that the log transform was the most appropriate. Due to the fact that there were a number of zeros for some of the dependent measures, the transform $X' = \log(X+1)$ was used for REP, MISS, INC and SIG while the transform X' = log(X) was used for RT. All subsequent data analyses were conducted using the log transformation of the data. Future mention of the transformed dependent measures will be preceded by the word "LOG" to denote that transformed data are being discussed.

The Multivariate Analysis of Variance (MANOVA) problem can be viewed as the process of finding the linear combinations of dependent variables which best separate the categories of the independent variables. Appendix B contains the complete sources of variation, the associated Sum of Squares (SS), degrees of freedom (df), and mean squares (MS) for the univariate tests for each of the dependent measures. Subsequent references to univariate statistics will be restricted to overall F values and their significance level.

Table 6 contains the multivariate and univariate summary results for all hypothesis tests. Results of the multivariate analysis for the test for the effects of aiding, illustrated in Figures 4, 5 and 6, indicate that there are significant differences among the aiding groups (df = 2,54; $p < .05$; Wilks' Criterion = .6158). Results of the univariate analyses revealed that this result may be attributed to differences in the following dependent measures: LOGRT ($F = 4.67$); SD ($F = 8.45$); and, LOGREP ($F = 4.65$). Results of Tukey's (A) HSD post hoc comparisons for these dependent measures indicate that there are no significant differences in terms of workload between the traditional manual mode and the intermediate aiding

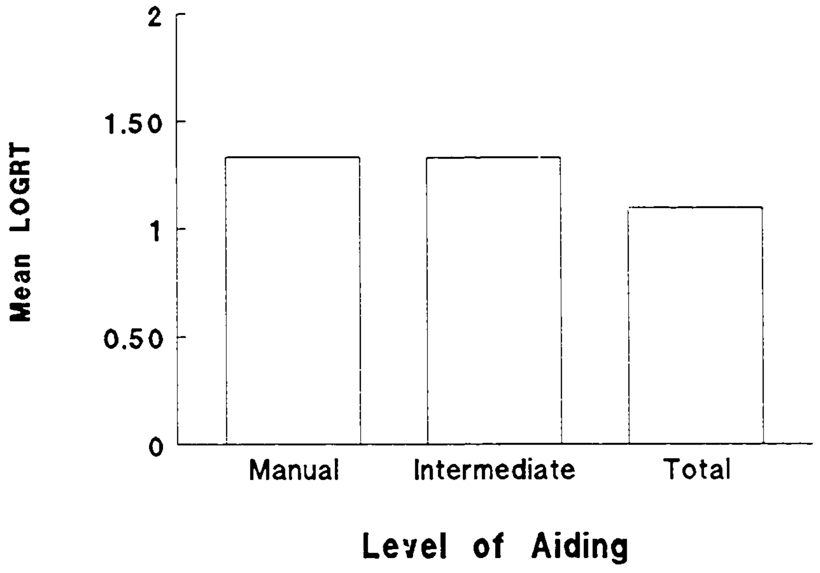


Figure 4. Aiding Main Effect Means: LOGRT

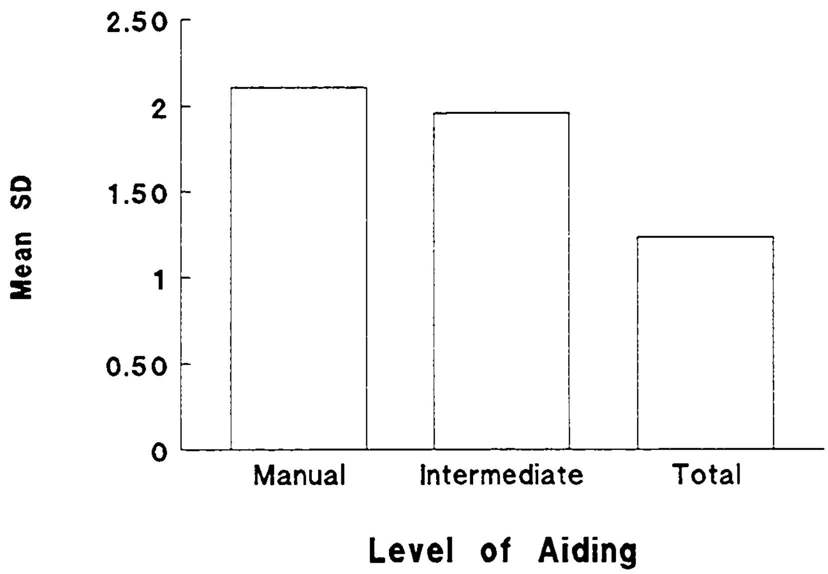


Figure 5. Aiding Main Effect Means: SD

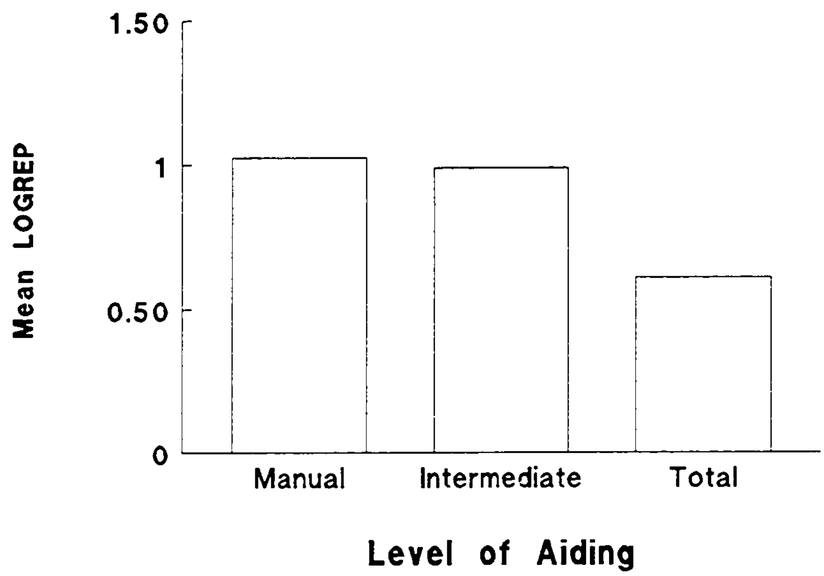


Figure 6. Aiding Main Effect Means: LOGREP

condition where advice was given to the individual and that to significantly reduce the amount of workload imposed on the individual, the task must be completely automated.

Significant differences (df = 1,54; $p < .05$; Wilks' Criterion = .7158) were also obtained for the test for differential effects on the dependent measures as a function of the primary task to which the subject was assigned. These results, which are illustrated in Figures 7, 8 and 9, indicate that there are differential effects across the two tasks on the following dependent measures: LOGRT (F = 10.25); SD (F = 5.71); and, LOGMISS (F = 5.46).

The results of the test to determine if learning occurred over sessions indicates that performance changed significantly over days (df = 2,108; $p < .05$; Wilks' Criterion = .6515). These results, which are illustrated in Figures 10, 11, 12, 13, 14 and 15, indicate that there are differential effects as a function of the length of practice on the following dependent measures: LOGRT (F = 26.80); SD (F = 25.19); LOGREP (F = 4.73); LOGMISS (F = 32.18); LOGINC (F = 7.59); and, LOGSIG (F = 15.93). Results of Tukey's (A) HSD post hoc comparison indicated that performance in Sessions 2 and 3 was significantly better than performance during the first session, but that there was no difference in performance between the latter two sessions.

No significant differences were noted for other statistical tests. The test for a differential effect of aiding as a function of primary task was not supported

Table 6. Secondary Task Multivariate and Univariate Summary Statistics

LOGRT	SD	LOGREP	LOGMISS	LOGINC	LOGSIG
Aiding Effect			Wilk's Criterion = .6158*		
4.67*	8.45*	4.65*	2.50	0.47	0.35
Primary Task Effect			Wilk's Criterion = .7158*		
10.25*	5.71*	3.67	5.46*	0.22	0.09
Session Effect			Wilk's Criterion = .6515*		
26.80*	25.19*	4.73*	32.18*	7.59*	15.93*
Primary Task by Aiding Effect			Wilk's Criterion = .7931		
1.01	1.71	0.16	3.39*	0.10	0.81
Aiding by Session Effect			Wilk's Criterion = .8135		
0.69	1.23	1.33	0.45	0.64	0.69
Primary Task by Session Effect			Wilk's Criterion = .9085		
0.51	0.63	1.32	0.85	0.85	2.59
Primary Task by Aiding by Session Effect			Wilk's Criterion = .8537		
0.84	1.12	0.28	2.70*	0.31	0.68

*denotes a *p* less than .05

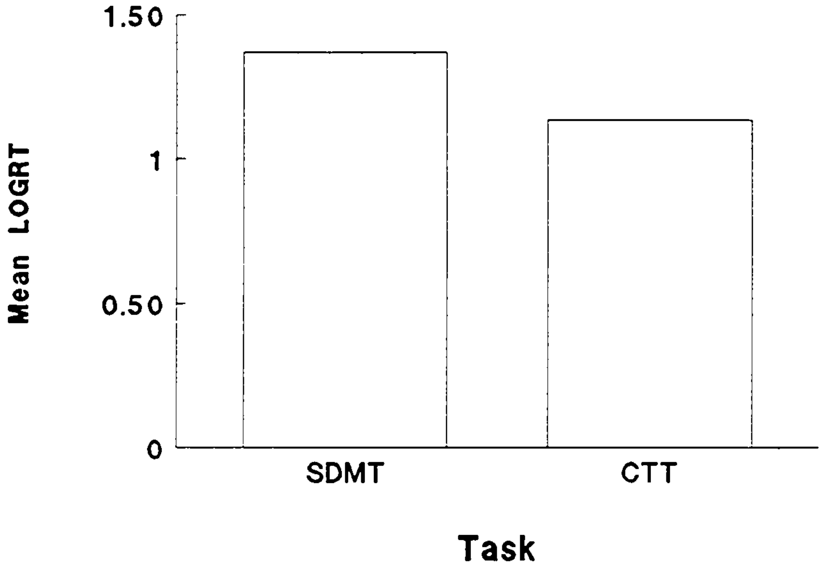


Figure 7. Task Main Effect Means: LOGRT

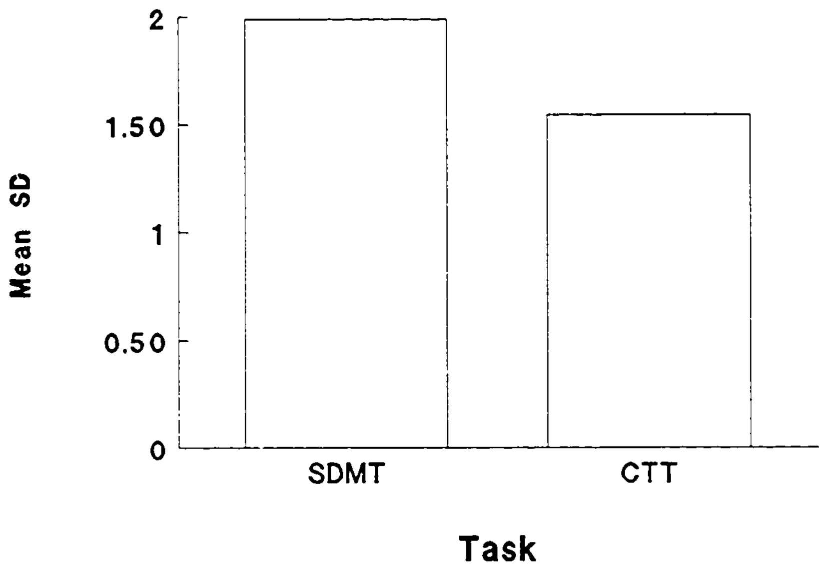


Figure 8. Task Main Effect Means: SD

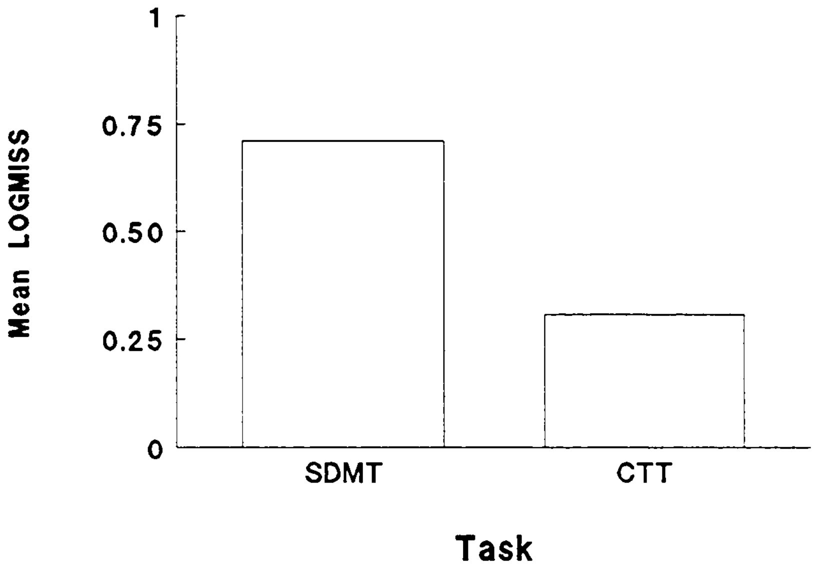


Figure 9. Task Main Effect Means: LOGMISS

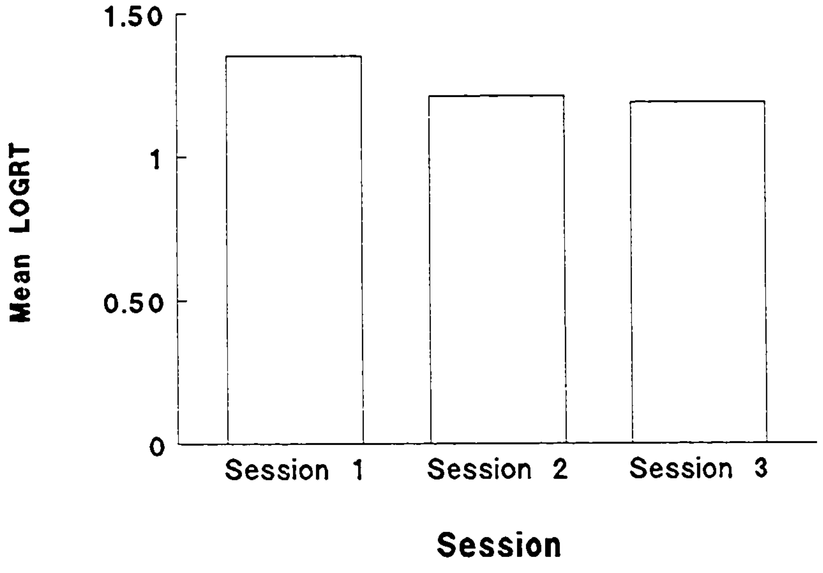


Figure 10. Session Main Effect Means: LOGRT

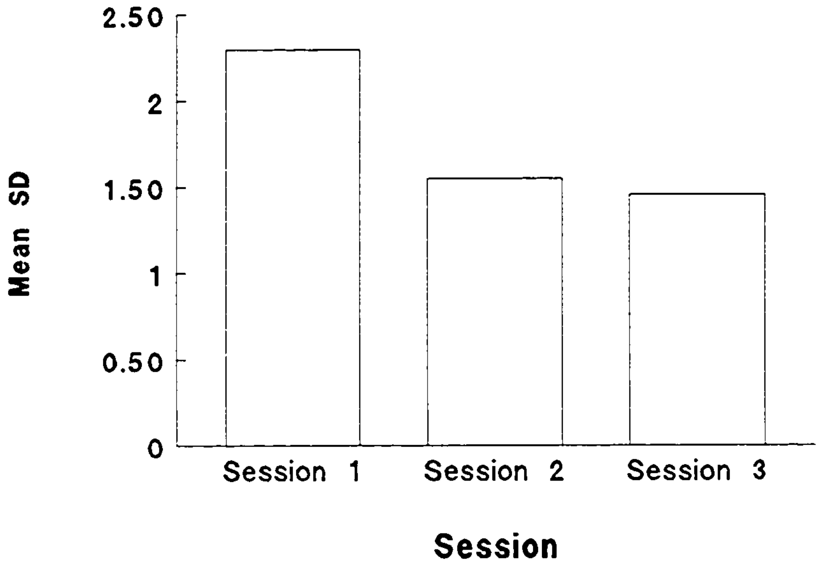


Figure 11. Session Main Effect Means: SD

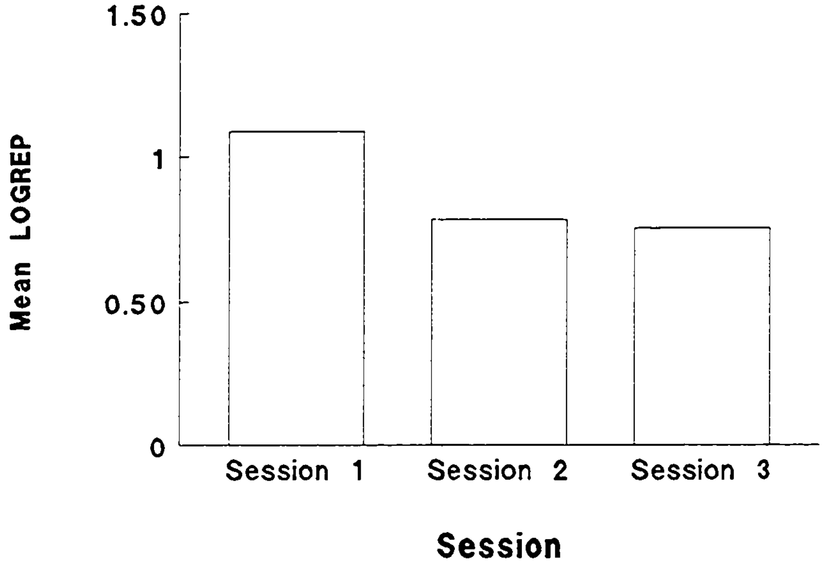


Figure 12. Session Main Effect Means: LOGREP

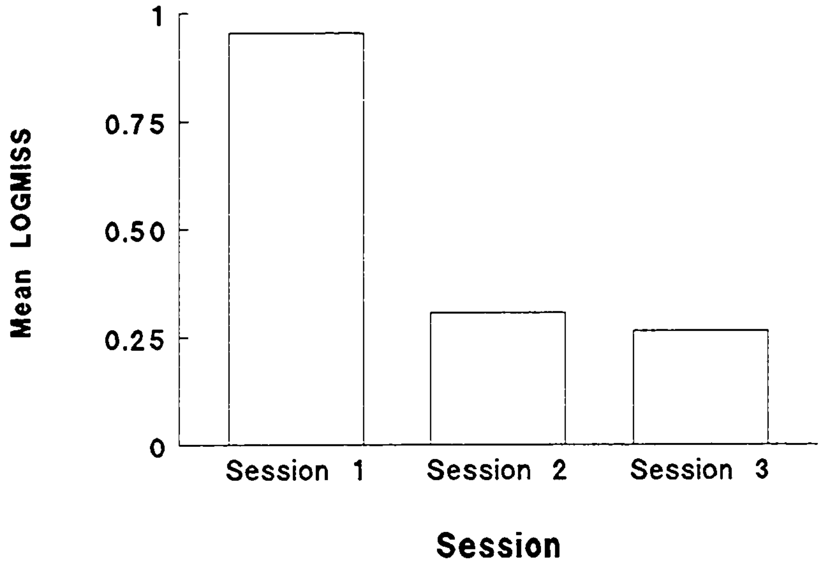


Figure 13. Session Main Effect Means: LOGMISS

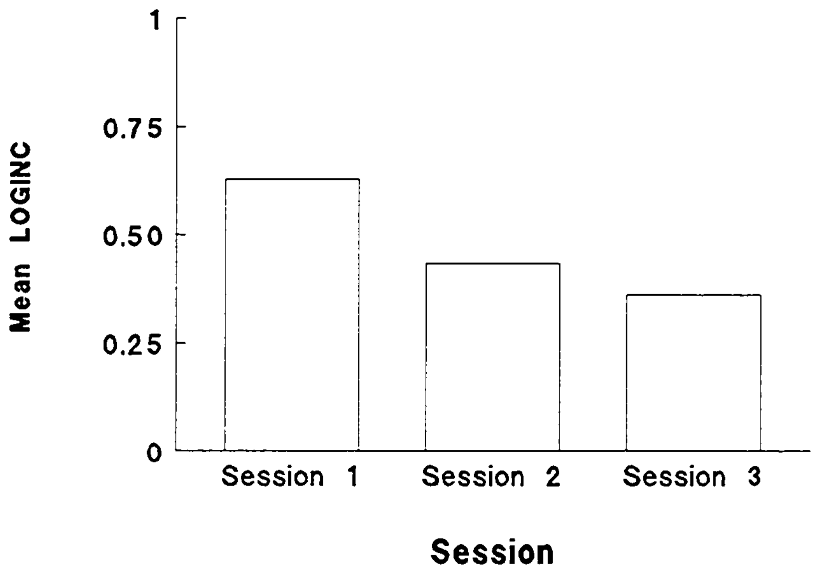


Figure 14. Session Main Effect Means: LOGINC

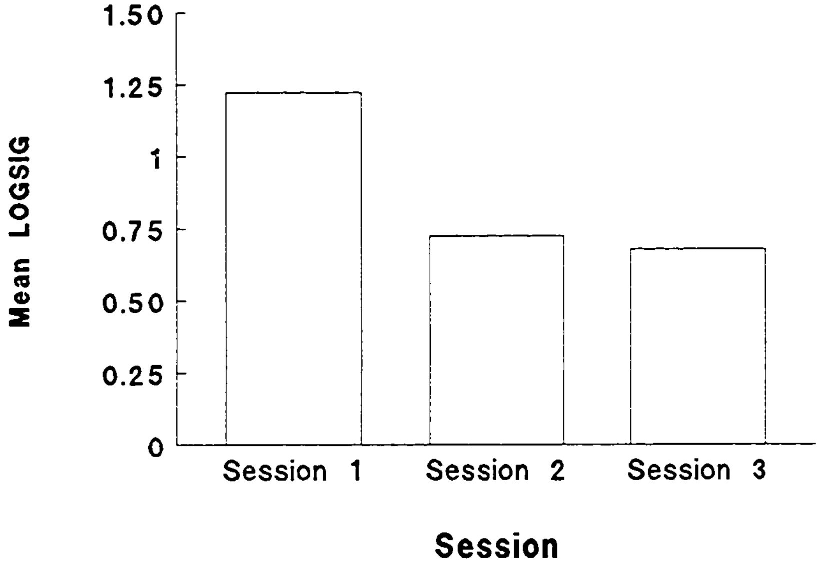


Figure 15. Session Main Effect Means: LOGSIG

(df = 2,54; p < .05; Wilks' Criterion = .7931), however, a significant univariate effect was obtained for LOGMISS (F = 3.39). A session effect as a function of the level of aiding employed was not obtained (df = 4,54; p < .05; Wilks' Criterion = .8135). In addition, the test that the session effect on secondary task errors was the same regardless of the primary task used was not supported (df = 2,54; p <.05; Wilks' Criterion = .9085) nor was the test for a differential effect of sessions on secondary task errors as a function of a particular combination of primary task and level of aiding (df = 4,54; p <.05; Wilks' Criterion = .8537).

DISCUSSION

The purposes of the present study were to investigate the effects of the addition of varying levels of automation on task workload requirements and to determine how this workload varied as a function of the type of task being automated. One hypothesis under investigation in the present study, which was that as the amount of cognitive workload increased, performance on a secondary task would be degraded, was not supported by the results of this investigation. The hypothesis that this decrease in performance would vary as a function of the primary task was supported.

Although significant differences were obtained when testing for the overall effect of aiding on performance, the obtained results were not consistent with that which might be predicted based on capacity theories. Capacity theories would predict that the increased mental operations required by the different levels of aiding would increase the competition for operator resources, thus resulting in a decrease in secondary task performance. Thus, the greatest decrement in performance should be noted in the totally automated condition which theoretically should impose the greatest demand for operator resources since the individual was required to continually verify that the system is performing the task correctly in addition to being prepared to perform the task manually if necessary. However, the data obtained in the present study indicated that a significant reduction in workload may be obtained by totally automating the task, as evidenced by significantly different (i.e., better) performance on LOGRT, SD and LOGREP. Differences on these measures were not noted for the other two conditions.

Significant differences were also obtained when testing for the hypothesis that the different primary tasks had no effects on the dependent measures. Dependent measures which were differentially affected by the primary task were LOGRT, SD and LOGMISS. In all cases better performance on these measures were noted with the CTT. This result may be explained in structural terms (cf., Keele, 1973), by assuming that the competing tasks impose simultaneous demands on specific perceptual or motor mechanisms. This assumption is consistent with a finding by Brown (1968) who noted that subsidiary tasks of interval production and random-number generation were affected differently by primary activities which involved a high rate of overt responses or a high rate of mental activity. The general rule is that the more similar the activity, the more they interfere with one another.

It is interesting to note that in addition to the more traditional dependent measures (i.e., reaction time, number of items missed, etc.) common to the verbal learning and retention literature (cf., Hall, 1971), there were other distinct types of responses made by the subjects in all of the conditions that are representative of errors which have been noted to occur in real world situations. Norman (1981), who analyzed motor errors (Fitts & Jones, 1961; Hurst, 1976; Reason, 1979) in an attempt to specify a theory of action, noted that errors could be classified as those which (1) result from errors in the formation of the intention; (2) result from the faulty activation of schemas; and, (3) result from the faulty triggering of active schemas. The distinctive behaviors noted in the present study for secondary task performance, i.e., REP, MISS, INC and SIG belong in the second category, which occur when schemas lose their activation before the appropriate time to control behavior has occurred and thereby results in the omission of components of the action sequence. These types of errors were noted to result from the normal decay and interference properties of primary memory. The fact that performance in the present situation resembles behavior in more complex situations lends credibility to any generalizations made from the results of the present study to performance in more complex situations considering the implementation of automation.

The results obtained from this study are disappointing in that significant differences were not obtained from tests of hypotheses concerning higher order interactions. There are two possible explanations why results such as these results might have been obtained. The first, and most obvious, is that the addition of less than total automation does not significantly impact the amount of workload which is required of the human operator. Given that this is a valid conclusion, the implications of this result for system designers is that the introduction of less than total automation should be restricted to more routine functions, for example, where there are large amounts of information that the operator must remember, and that tasks or functions should be totally automated in those situations where reduction of operator workload is the primary justification for introducing automation.

Another possible explanation for the results obtained in the present study is that the secondary task methodology may not be sufficiently sensitive to the changes in cognitive workload imposed by the various gradations of automation used in the present study. The primary measures of interest in the present study were the decision processes which resulted in certain actions or inaction and the factors which influenced these decisions (e.g., the state of the system variables, what information is available, preceding events, the operator's knowledge and mental model of the system). Since these performance measures differ in nature and are more covert than more traditional performance measures, which are easily manipulated and measured, more subjective methodologies, such as verbal protocols (cf., Bainbridge, 1974; Ericsson & Simon, 1980) and other more formal psychometric techniques such as multiattribute utility assessment or multidimensional scaling might be more appropriate to research of this type.

Given the high degree of user acceptance of automation (Kibbe & DeVere, 1987), it seems inevitable that automation will be implemented in many complex task situations. This points to the need for a comprehensive data base which details the situations under which and the extent to which automation is needed is essential in order that HF is responsive to the issues attendant to the effective design of these systems. Ideally, a program of research to address issues of this type should proceed from an analysis of the "total picture" in order to avoid the problems inherent in others areas of HFE research where it is difficult to formulate general principles of human performance. The stress literature is one example where research studies typically have been done to address specific issues in regard to specific situations rather proceeding from an analysis of the total picture. By assessing all the different functional types of tasks which might be candidates for future aiding as well as the various types of aiding which might be available, and tailoring the research along these lines, it might be possible to conduct research from which general principles of human performance and guidelines for the implementation of automation in complex systems might be developed.

One major purpose of this study was to identify areas of needed research concerning the performance implications of the introduction of task automation. When such research is completed, the full significance of the results obtained in the present study perhaps may be assessed.

REFERENCES

Al-Awar, J. A., Chapanis, A. & Ford, W.R. (1981). Tutorials for the first-time computer user. IEEE Transactions on Professional Communication, PC-24, 30-37.

Allen, R.R. (Ed.). (1984). Computer text editing [Special Issue]. Human Factors, 26(4).

Alluisi, E.A. (1967). Methodology in the use of synthetic tasks to assess complex performance. Human Factors, 9, 375-384.

Alluisi, E.A. (1969). Sustained Performance. In: E.A. Bilodeau & I. Bilodeau (Eds.), Principles of skill acquisition (pp. 59-101). New York: Academic Press.

Alter, S. (1980). Decision support systems: Current practices and continuing challenges. Reading: Addison-Wesley.

Anderson, J. & Bower, G. (1973). Human associative memory. Washington, DC: Winston.

Bainbridge, L. (1982). Ironies of Automation. In: G. Johannsen & J.E. Rijnsdorp (Eds.), Analysis, design and evaluation of man-machine systems (pp. 125-135). Oxford: Pergamon Press.

Bainbridge, L. (1974). The representation of working storage and its use in the organisation of behavior. In: W.T. Singleton & P. Spurgeon (Eds.), Measurement of human resources (pp. 165-183). London: Taylor and Francis.

Bahrick, H.P., Noble, M. & Fitts, P.M. (1954). Extra-task performance as a measure of learning a primary task. Journal of Experimental Psychology, 48, 298-302.

Barnard, P.J., Hammond, N.V., Morton, J. & Long, J.B. (1981). Consistency and compatibility in human-computer dialogue. International Journal of Man-Machine Studies, 15, 87-123.

Barr, A. & Feigenbaum, E.A. (Eds.). (1981). The handbook of artificial intelligence (Volume 1). Stanford: HeurisTech Press.

Barr, A. & Feigenbaum, E.A. (Eds.). (1982a). The handbook of artificial intelligence (Volume 2). Stanford: HeurisTech Press.

Barr, A. & E.A. Feigenbaum (Eds.). (1982b). The handbook of artificial intelligence (Volume 3). Stanford: Heuristech Press.

Bennett, J.L. (1979). Incorporating usability into system design. Design '79 Symposium. Monterey, CA.

Berry, D.C. & Broadbent, D.E. (1987). Expert systems and the man-machine interface: Part II. The user interface. Expert Systems, 4, 18-27.

Boehm-Davis, D.A., Curry, R.E., Weiner, E.L. & Harrison, R.L. (1981). Human factors of flight deck automation - NASA/Industry workshop. (NASA TM 81260). Moffett Field, CA: National Aeronautics and Space Administration.

Bilodeau, E.A. & I. Bilodeau (Eds.). (1969). Principles of skill acquisition. New York: Academic Press.

Borg, G. (1978a). Subjective aspects of physical work. Ergonomics, 21, 215-220.

Borg, G. (1978b). Subjective effort in relation to physical performance and working capacity. In: H. Pick (Ed.), Psychology from research to practice (pp. 333-361). New York: Plenum.

Broadbent, D.E. (1957). A mechanical model for human attention and immediate memory. Psychological Review, 64, 205-215.

Broadbent, D.E. (1958). Perception and communication. London: Pergamon.

Broadbent, D.E. (1971). Decision and stress. New York: Academic Press.

Brown, E.L., Stone, G. & Pearce, W.E. (1975). Improved cockpits through flight crew workload measurement. Paper presented at the Second Advanced Aircrew Display Symposium. Patuxent River, MD.: Naval Air Test Center.

Brown, I.D. (1964). The measurement of perceptual load and reserve capacity. Transactions of the Association of Industrial Medical Officers, 14, 44-49.

Brown, I.D. (1968). Criticisms of time-sharing techniques for the measurement of perceptual-motor difficulty. Proceedings of the XVI International Congress of Applied Psychology (pp. 322-328). Amsterdam: Swets & Zeitlinger.

Card, S.K., English, W.K. & Burr, B.J. (1978). Evaluation of mouse, rate-controlled isometric joystick, step keys and text keys for text selection on a CRT. Ergonomics, 21, 601-631.

Carter, J.A. (1986). A taxonomy of user-oriented functions. International Journal of Man-Machine Studies, 24, 195-292.

Chapanis, A. (1982). Humanizing computers. Proceedings of the ITT Europe Human Factors Symposium (pp. 20-53). London: ITT Europe Engineering Support Centre.

Chiles, W.D. (1982). Workload, task, and situational factors as modifiers of complex human performance. In: E.A. Alluisi & E.A. Fleishman (Eds.), Human performance and productivity: Stress and performance effectiveness (pp. 11-56). Hillsdale: Lawrence Earlbaum Associates.

Chiles, W.D. & Alluisi, E.A. (1979). On the specification of operator or occupational workload with performance measurement methods. Human Factors, 21(5), 515-528.

Chiles, W.D., Alluisi, E.A. & Adams, O.S. (1968). Work schedules and performance during confinement. Human Factors, 10, 143-196.

Clement, W.F., McRuer, D.T. & Klein, R.H. (1971). Systematic manual control display design. In: Advisory Group for Aerospace Research and Development (Ed.), Proceedings of the AGARD Conference on Guidance and Control Displays (pp. 6-1 - 6-10). London: Harford House.

Cohill, A.M. & Williges, R.C. (1985). Retrieval of HELP information for novice users of interactive computer systems. Human Factors, 27, 335-344.

Cooper, G.E. & Harper, R.P. (1969). The use of pilot rating in the evaluation of aircraft handling qualities (NASA TN D-5153). Washington, DC: National Aeronautics and Space Administration.

Craik, K.H.W. (1948). Theory of the human operator in control systems. British Journal of Psychology, 38, 142-144.

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

Cumming, F.G. & Corkindale, K.G. (1967). Physiological and psychological measurements of pilot workload (TM HFG 101). Farnborough: Royal Aircraft Establishment.

Davis, R. (1956). The limit of the psychological refractory period. Quarterly Journal of Experimental Psychology, 8, 13-38.

Davis, R. (1957). The human operator as a single channel information system. Quarterly Journal of Experimental Psychology, 9, 119-129.

Davis, R. (1959). The role of attention in the psychological refractory period. Quarterly Journal of Experimental Psychology, 11, 211-220.

Davis, R. (1962). Choice reaction times and the theory of intermittency in human performance. Quarterly Journal of Experimental Psychology, 14, 157-166.

Davis, R. & King, J.J. (1977). An overview of production systems. In: E. Elcock & D. Michie (Eds.), Machine intelligence 8 (pp. 300-332). Chichester: Ellis Horwood.

Davis, R. & Lenat, D. (1982). Knowledge-based systems in artificial intelligence. New York: McGraw-Hill.

Defense Advanced Research Projects Agency (1983). Strategic Computing. Washington, DC.

DeKleer, J. & Brown, J.S. (1983). Assumptions and ambiguities in mechanistic mental models. In: D. Gentner & A.L.Stevens (Eds.), Mental models (pp. 155-190). Hillsdale: Lawrence Erlbaum.

Department of Defense (1977a). Major System Acquisition. (DOD Directive 5000.1). Washington, DC: Author.

Department of Defense (1977b). Major System Acquisition Process. (DOD Directive 5000.2). Washington, DC: Author.

Department of Defense (1977c). Test and Evaluation. (DOD Directive 5000.3). Washington, DC: Author.

Duncan, J. (1980). The demonstration of capacity limitation. Cognitive Psychology, 12, 75-96.

Eason, K.D. (1974). The manager as a computer user. Applied Ergonomics, 5, 9-14.

Ericsson, K.A. & Simon, H.A. (1980). Verbal reports as data. Psychological Review, 87, 215-251.

Eggemeier, F.T. (1981). Current issues in subjective assessment of workload. Proceedings of the Human Factors Society 25th Annual Meeting, 25, 513-517.

Feigenbaum, E.A. (1982). Knowledge engineering for the 80's. Palo Alto: Stanford University Computer Sciences Department.

Fitts, P.M. & Jones, R.E. (1961). Analysis of factors contributing to 460 "pilot-error" experiences in operating aircraft controls. In: W.J. Sinaiko (Ed.), Selected papers on human factors in the design and use of control systems. New York: Dover.

Funt, B.V. (1976). WHISPER: A computer implementation using analogs in reasoning (RN 76-09). Vancouver, British Columbia: University of British Columbia.

Gaines, B.R. & Shaw, M.L.G. (1986). From timesharing to the sixth generation: The development of human-computer interaction. Part I. International Journal of Man-Machine Studies, 24, 1-27.

Gartner, W.B. & Murphy, M.R. (1976). Pilot workload and fatigue: A critical survey of concepts and assessment techniques (NASA TN D-8365). Washington, DC: National Aeronautics and Space Administration.

Garvey, W.D. & Henson, J.B. (1958). Interactions between display gain and task-induced stress in manual tracking systems (NRL RN 5204). Washington, DC: Naval Research Laboratory.

Gentner, D. & Stevens, A.L. (1983). Mental models. Hillsdale: Lawrence Erlbaum.

Gevarter, W.B. (1982). An overview of computer vision (NSBIR 82-2582). Washington, DC: National Bureau of Standards.

Gevarter, W.B. (1983a). An overview of artificial intelligence and robotics. Part B: Applications (NASA TM 85838). Washington, DC: National Aeronautics and Space Administration.

Gevarter, W.B. (1983b). An overview of computer-based natural language processing (NASA TM 85635). Washington, DC: National Aeronautics and Space Administration.

Gordon, D.E. (1985). Manual and computer-aided sequential diagnostic inference (AFHRL TP 84-51). Wright Patterson Air Force Base, Ohio: Logistics and Human Factors Division.

Grandjean, E. & Vigliani, E. (Eds.). (1980). Ergonomic aspects of visual display terminals. London: Taylor and Francis.

Greenstein, J.S. & Revesman, M.E. (1986). Human-computer communication for dynamic task allocation. In: G. Salvendy (Ed.), Advances in human factors/ergonomics (pp. 319-366). Amsterdam: Elsevier.

Halasz, F.G. & Moran, T.P. (1983). Mental models and problem solving in using a calculator. In: A. Janda (Ed.), Human factors in computing systems (pp. 212-216). Boston, MA: Association for Computing Machinery, Inc.

Hall, J.F. (1971). Verbal learning and retention. Philadelphia: J.B. Lippincott.

Hart, S.G. & Sheridan, T.B. (1984). Pilot workload, performance and aircraft control automation. Advisory Group for Aerospace Research and Development (Ed.), Proceedings of the AGARD Symposium on Human Factors Contributions to High Performance Aircraft. London: Harford House.

Hartman, B.O. & McKenzie, R.E. (Eds.). (1979). Survey of methods to assess workload (AGARD-AG-246). London: Harford House.

Hayes-Roth, F., Waterman, D.A. & Lenat, D.B. (Eds.). (1983). Building Expert Systems. Reading: Addison-Wesley.

Hellyer, S. (1962). Supplementary report: Frequency of stimulus presentation and short-term decrement in recall. Journal of Experimental Psychology, 64, 650.

Hewitt, C. (1975). How to use what you know. International Joint Conferences on Artificial Intelligence, 4, 198-198.

Hick, W.E. (1949). Reaction time for the amendment of a response. Quarterly Journal of Experimental Psychology, 1, 175-179.

Hochberg, J.E. (1970). Attention, organization and consciousness. In: D.I. Mostofsky (Ed.), Attention: Contemporary theory and analysis (pp. 99-124). New York: Appleton-Century-Crofts.

Holding, D.H. (Ed.). (1981). Human Skills. Chichester: Wiley.

Hollnagel, E. & Woods, D. (1983). Cognitive systems engineering: New wine in new bottles. International Journal of Man-Machine Studies, 18, 583-600.

Hurst, E. (Ed.). (1976). Pilot error: A professional study of contributory factors. London: Butler & Tanner.

Hutchinson, R.D. (1981). New horizons for human factors in design. New York: McGraw-Hill.

Jahns, D.W. (1973). Operator workload: What is it and how should it be measured? In: K.D. Gross & J.J. McGrath (Eds.), Crew system design (pp. 281-288). Santa Barbara, CA.: Anacapa Sciences, Inc.

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

Jerison, H.J. & Pickett, R.M. (1963). Vigilance: A review and re-evaluation. Human Factors, 5, 211-238.

Johannsen, G. (1979). Workload and workload measurement. In: N. Moray (Ed.), Mental Workload: Its theory and measurement (pp. 3-11). New York: Plenum Press.

Johassen, D.H. (Ed.). (1981). The technology of text. Englewood Cliffs: Educational Technology Publications.

Kahneman, D. (1973). Attention and effort. New York: Prentice-Hall.

Kantowitz, B.H. & Knight, J. (1976). On experimenter-limited processes. Psychological Review, 83, 502-507.

Keele, S.W. (1973). Attention and human performance. Pacific Palisades, CA.: Goodyear.

Kelley, C.R. (1969). The measurement of tracking proficiency. Human Factors, 11, 43-64.

Kelley, C.R. & Wargo, M.J. (1967). Cross-adaptive operator loading tasks. Human Factors, 9, 395-404.

Kibbe, M.J. & DeVere, R.P. (1987). Aircrew opinions on information displays and automation for electronic warfare systems (NWC-TP-6782). China Lake, CA: Naval Weapons Center.

Kidd, A.L. & Cooper, M.B. (1985). Man-machine interface issues in the construction and use of an expert system. International Journal of Man-Machine Studies, 22, 91-102.

Knowles, W.B. (1963). Operator loading tasks. Human Factors, 5, 155-161.

Lane, D.M. (1982). Limited capacity, attention allocation, and productivity. In: W.C. Howell and E. A. Fleishman (Eds.), Human performance and productivity: Information processing and decision making (pp.121-156). Hillsdale: Lawrence Earlbaum.

Lehnert, W.G. & Ringle, M.H. (Eds.) (1982). Strategies for natural language processing. Hillsdale, NJ: Lawrence Erlbaum.

Manna, Z. (1973). Introduction to the mathematical theory of computation. New York: McGraw-Hill.

Marr, D.C. (1982). Vision. San Francisco: W.H. Freeman.

Martin, J. (1973). Design of man-computer dialogues. Englewood Cliffs: Prentice-Hall.

McCormick, E.J. & Sanders, M.S. (1982). Human Factors in engineering and design. New York: McGraw-Hill.

Meister, D. (1976). Behavioral foundations of system development. New York: Wiley.

Meister, D. (1982). The role of Human Factors in system development. Applied Ergonomics, 13, 119-124.

Mischkoff, H.C. (1985). Understanding artificial intelligence. Dallas: Texas Instruments, Inc.

Moran, T. (1981). An applied psychology of the user. Computing Surveys, 13, 1-11.

Moray, N. (1967). Where is capacity limited? A survey and a model. Acta Psychologica, 27, 84-93.

Moray, N. (1979). Models and measures of mental workload. In: N. Moray (Ed.), Mental Workload: Its theory and measurement (pp. 13-21). New York: Plenum Press,

Moray, N. (1981). Feedback and the control of skilled behaviour. In: D.H. Holding (Ed.), Human skills (pp. 15-39). Chichester: Wiley.

Moray, N. (1982). Subjective mental workload. Human Factors, 24(1), 25-40.

Morgan, B.B. Jr. & Alluisi, E.A. (1972). Synthetic work: Methodology for assessment of human performance. Perceptual and Motor Skills, 35, 835-845.

National Research Council (1983). Research needs for human factors. Washington, DC: National Academy Press.

National Research Council (1984). Research and Modeling of Supervisory Control Behavior. Washington, DC: National Academy Press.

Navon, D. & Gopher, D. (1979). On the economy of the human processing system. Psychological Review, 86, 217-25.

Neisser, U. (1970). Cognitive psychology. New York: Appleton-Century-Crofts.

Newell, A. & Rosenbloom, P.S. (1981). Mechanisms of skill acquisition and the law of practice. In: J.R. Anderson (Ed.), Cognitive skills and their acquisition (pp. 1-55). Hillsdale: Lawrence Erlbaum.

Norman, D.A. (1968). Towards a theory of memory and attention. Psychological Review, 75, 522-536.

Norman, D.A. (1981). Categorization of action slips. Psychological Review, 88(1), 1-15.

Norman, D.A. & Bobrow, D. (1975). On data-limited and resource limited processes. Cognitive Psychology, 7, 44-64.

Norman, D.A. & Rumelhart, D.E. (Eds.). (1975). Explorations in cognition. San Francisco: W.H. Freeman.

North, R. (1977). Task components and demands as factors in dual task performance (TR-ARL-77-4). Champaign, IL: University of Illinois at Urbana-Champaign Aviation Research Laboratory.

Noyes, J. (1983). The QWERTY keyboard: a review. International Journal of Man-Machine Studies, 18, 265-281.

Odgen, G.D., Levine, J.M. & Eisner, E.J. (1979). Measurement of workload by secondary tasks. Human Factors, 21(5), 529-548.

Parsons, H. McIllvaine. (1985). Automation and the individual: Comprehensive and comparative views. Human Factors, 27, 99-112.

Peterson, L.R. & Peterson, M.J. (1959). Short term retention of individual verbal items. Journal of Experimental Psychology, 59, 193-198.

Pew, R.W. (1979). Secondary tasks and workload measurement. In: N. Moray (Ed.), Mental workload: Its theory and measurement (pp. 23-28). New York: Plenum Press.

Posner, M.J. & Boies, S.J. (1971). Components of attention. Psychological Review, 78, 391-408

Price, H.E. (1980a). Department of Defense and service requirements for human factors R & D in the military system acquisition process (RN 80-23). Alexandria, VA.: U.S. Army Research Institute.

Price, H.E. (1980b). The contribution of human factors in military systems development: Methodological considerations (TR 476). Alexandria, VA: U.S. Army Research Institute.

Price, H.E. (1985). The allocation of function in systems. Human Factors, 27, 33-46.

Ramsay, H.R. & Atwood, M.E. (1979). Human factors in computer systems: A review of the literature (TR SAI-79-111-DEN). Englewood, CO: Science Applications, Inc.

Reason, J.T. (1977). Skill and error in everyday life. In: M. Howe (Ed.), Adult learning. London: Wiley.

Reid, G.B., Shingledecker, C.A. & Eggemeier, F.T. (1981). Applications of conjoint measurement to workload scale development. Proceedings of the Human Factors Society 25th Annual Meeting, 25, 522-526.

Rich, E. (1979). User modeling through stereotypes. Cognitive Science, 3, 329-354.

Rich, E. (1983). Users are individuals: Individualizing user models. International Journal of Man-Machine Studies, 18, 199-214.

Roscoe, A.H. (Ed.). (1978). Assessing pilot workload (AGARD CP-216). London: Harford House.

Rolfe, J.M. (1971). The secondary task as a measure of mental load. In: W.T. Singleton, J.G. Fox & D. Whitfield (Eds.), Measurement of men at work (pp. 135-148). London: Taylor and Francis.

Rolfe, J.M. & Lindsey, S.J.E. (1973). Flight deck environment and pilot workload: Biological measures of workload. Applied Ergonomics, 4, 199-206.

Rouse, W.B. (1977). Human-computer interaction in multi-task situations. IEEE Transactions on Systems, Man and Cybernetics, SMC-7(5), 384-392.

Rouse, W.B. (1981). Pilot interaction with automated decision making systems (NASA CR 164-729). Washington, DC: National Aeronautics and Space Adminstration.

Rouse, W.B. & Rouse, S. (1983). A framework for research in adaptive decision aids (AFAMRL-TR-83-082). Wright-Patterson AFB: Air Force Aerospace Medical Research Laboratory.

Salvendy, G. (Ed.). (1987). Handbook of human factors. New York: Wiley.

Schneider, W. & Shriffrin, R.M. (1977). Controlled and automatic human information processing: I. Detection, search and attention. Psychological Review, 84(1), 1-66.

Self, J.A. (1977). Concept teaching. Artificial Intelligence, 9, 197-221.

Selye, H. (1956). The stress of life. New York: McGraw-Hill.

Sharit, J. (1985). Supervisory control of a flexible manufacturing system. Human Factors, 27, 47-60.

Sheridan, T.B. (1976). Towards a general model of supervisory control. In: T.B. Sheridan & G. Johannsen (Eds.), Monitoring behavior and supervisory control (pp. 271-281). New York: Plenum Press.

Sheridan, T.B. (1984). Supervisory control of remote manipulators, vehicles and dynamic processes. In: W.B. Rouse (Ed.), Advances in man-machine research (pp. 49-137). New York: JAI Press.

Sheridan, T.B. & Johannsen, G. (Eds.). (1976). Monitoring Behavior and Supervisory Control. New York: Plenum Press.

Sheridan, T.B. & Stassen, H.G. (1979). Definitions, models and measures of human workload. In N. Moray (Ed.), Mental workload: Its theory and measurement (pp. 219-232). New York: Plenum Press.

Shneiderman, B. (1980). Software psychology: Human factors in computer and information systems. Cambridge: Winthrop Publishers.

Shriffrin, R.M. & Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending and a general theory. Psychological Review, 84(2), 127-190.

Smith, S.L. (1981). The usability of software: Design guidelines for the user-system interface. Paper presented at the ITT Symposium on Human Factors and the Usability of Software, Sheldon, CONN.

Singleton. W.T. (1971). Prologue to Section 1: Man. In: W.T. Singleton, J.G. Fox & D. Whitfield (Eds.), Measurement of man at work (pp. 15-16). London: Taylor and Francis.

Sloman, A. (1975). Afterthoughts on analogical representations. Workshops on Theoretical Issues in Natural Language Processing, 1, 178-182.

Stewart, T.F.M. (1974). Ergonomic aspects of man-computer problem solving. Applied Ergonomics, 5, 209-212.

Swedish, W.J. (1983). Evolution of advanced ATC automation functions (Mitre Working Paper Report 83W149). McLean, VA: The Mitre Corporation.

Tannas, L.E. (1985). Flat-panel displays and CRTs. New York: Van Nostrand Reinhold.

Thackeray, R.I. (1980). Boredom and monotony as a consequence of automation: A consideration of the evidence relating boredom and monotony to stress (Report No. FAA-AM-80-1). Washington, DC: Federal Aviation Administration.

Triesman, A.M. (1969). Strategies and models of selective attention. Psychological Review, 76(3), 282-299.

Triesman, A.M. & Davies, A. (1973). Divided attention to ear and eye. In: S. Kornblum (Ed.), Attention and performance IV. London: Academic Press.

VanCott, H.P. & Kinkade, R.G. (Eds.). (1972). Human engineering guide to equipment design. Washington, DC: U.S. Government Printing Office.

Vaughan, W.S. & Mavor, A.S. (1972). Behavioral characteristics of men in performance of some decision-making task components. Ergonomics, 15(3), 267-277.

Warm, J.S. (1984). Sustained attention in human performance. Chichester: Wiley.

Waterman, D.A. & Hayes-Roth, F. (Eds.). (1978). Pattern directed inference systems. New York: Academic Press.

Welford, A.T. (1952). The psychological refractory period and the time of high speed performance -- a review and a theory. British Journal of Psychology, 43, 2-19.

Welford, A.T. (1953). The psychologist's problem in measuring fatigue. In: W.F. Floyd & A.T. Welford (Eds.), Symposium on fatigue (pp. 183-191). London: H.K. Lewis and Co.

Welford, A.T. (1968). Fundamentals of skill. London: Metheun.

Welford, A.T. (1973). Stress and performance. Ergonomics, 16, 567-580.

White, R.T. (1971). Task analysis methods: Review and development of techniques for analyzing mental workload in multiple task situations (MDC J 5291). St. Louis, MO: McDonnell Douglas Corporation.

Wickens, C.D. (1979). Workload, stress and secondary tasks. In: N. Moray (Ed.), Mental workload: Its theory and measurement (pp. 79-99). New York: Plenum Press.

Wickens, C.D. (1980). The structure of attentional resources. In: R. Nickerson & R. Pew (Eds.), Attention and performance VIII. Hillsdale: Lawrence Earlbaum.

Wickens, C.D., Montford, S.J. & Schreiner, W. (1981). Multiple resources, task-hemispheric integrity and individual differences in time-sharing. Human Factors, 23, 211-230.

Wierwille, W. (1979). Physiological measures of aircrew mental workload. Human Factors, 21(5), 575-593.

Wierwille, W. & Williges, R.C. (1978). Survey and analysis of operator workload assessment techniques. (Report No. S-78-101). Blacksburg, VA: Systemetrics, Inc.

Williges, B.H. & Williges, R.C. (1984). Dialogue design considerations for interactive computer systems. In: F. Muckler (Ed.), Human factors review: 1984 (pp. 167-208). Santa Monica, CA: Human Factors Society.

Williges, R.C. (1982). Applying the human information processing approach to human/computer interactions. In: W.C. Howell & E.A. Fleishman (Eds.), Human performance and productivity: Information processing and decision making (pp. 83-119). Hillsdale: Lawrence Erlbaum.

Williges, R.C. & Ehrich, R.W. (1984). Multidisciplinary research in human-computer dialogue design. In: G. Salvendy (Ed.), Advances in human factors/ergonomics II (pp. 1-9). Amsterdam: Elsevier.

Williges, R.C. & Wierwille, W. (1979). Behavioral measures of aircrew mental workload. Human Factors, 21(5), 549-574.

Williges, R.C. & Williges, B.H. (1982). Human-computer dialogue design considerations. In: G. Johannsen & J.E. Rijnsdorp (Eds.), Analysis, design and evaluation of man-machine systems (pp. 239-246). Oxford: Pergamon Press.

Winograd, T. (1972). Understanding natural language. New York: Academic Press.

Winston, P.H. (1977). Artificial intelligence. Reading: Addison-Wesley.

Winston, P.H. (1982). Artificial Intelligence. Cambridge: The MIT Press.

Woods, D.D. (1985). Knowledge based development of graphic display systems. Proceedings of the Human Factors Society 29th Annual Meeting, 29, 325-329.

Woods, D.D. (1986). Paradigms for intelligent decision support. In: E. Hollnagel (Ed.), Intelligent Decision Support in Process Environments (pp. 153-173). Heidelberg: Springer-Verlag.

Zachary, W. (1986). A cognitively based functional taxonomy of decision support techniques. Human-Computer Interaction, 2(1), 25-63.

Zachary, W., Wherry, R.J. & Glenn, F.A. (1981). Framework for decision augmentation system design: Structure and organization (TR1400.11B). Willow Grove, PA: Analytics.

Zakay, D. (1982). Reliability of information as a potential threat to the acceptability of decision support systems. IEEE Transactions on Systems, Man and Cybernetics, SMC-12(4), 518-520.

APPENDICES

Appendix A. Subject Instructions

C1 (Primary Task - Decision Making Level of Aiding - Manual)

During the next hour you will be asked to perform two tasks -- a decision making and long term memory task. In the decision making task you will be making one of three responses to various symbols which will be presented in the box at the lower left portion of the screen. The symbols you will be presented with are:

- a "#" to which you will respond "C";
- an "*" to which you will respond "Z"; and,
- an "@" to which you will respond "X".

In addition, you will also be performing a long term memory task. At random intervals throughout the session you will see the following message in the box at the lower right portion of the screen: "ENTER CURRENT VALUE". Your task will be to mentally subtract three from the previous number and input this resulting value into the computer using the number keys at the right of the keyboard. The experimenter will tell you which number you should start subtracting from at the beginning of the session.

It is important to keep one thing in mind: During the session the decision making task is of primary importance AND SHOULD BE PERFORMED BEFORE THE LONG TERM MEMORY TASK, should they occur together.

Please take this opportunity to go over these instructions with the experimenter and ask any questions you might have.

C2 (Primary Task - Decision Making Level of Aiding - Advice)

During the next hour you will be asked to perform two tasks -- a decision making and long term memory task. In the decision making task you will be making one of three responses to various symbols which will be presented in the box at the lower left portion of the screen. The symbols you will be presented with are:

- a "#" to which you will respond "C";
- an "*" to which you will respond "Z"; and,
- an "@" to which you will respond "X".

In order to help you make this decision, when the symbols are presented in the lower left hand box you will also be receiving advice on how to respond to the task. You will need to verify the advice you receive (it may not necessarily be correct so it is important to check) and then to input the proper response into the computer.

In addition, you will also be performing a long term memory task. At random intervals throughout the session you will see the following message in the box at the lower right portion of the screen: "ENTER CURRENT VALUE". Your task will be to mentally subtract three from the previous number and input this resulting value into the computer using the number keys at the right of the keyboard. The experimenter will tell you which number you should start subtracting from at the beginning of the session.

It is important to keep one thing in mind: During the session the decision making task is of primary importance AND SHOULD BE PERFORMED BEFORE THE LONG TERM MEMORY TASK, should they occur together.

Please take this opportunity to go over these instructions with the experimenter and ask any questions you might have.

C3 (Primary Task - Decision Making Level of Aiding - Total)

During the next hour you will be asked to perform two tasks -- a decision making and long term memory task. In the decision making task you will be making one of three responses to various symbols which will be presented in the box at the lower left portion of the screen. The symbols you will be presented with are:

- a "#" to which you will respond "C";
- an "*" to which you will respond "Z"; and,
- an "@" to which you will respond "X".

You will be receiving some help in performing this task. When the symbols are presented the computer will also make a response which may or may not be correct. Your task in this session will be to verify that the response which was made by the computer was correct (it may not necessarily be correct so it is important to keep checking that it is done properly). If you have determined that it is correct, then press the space bar to continue, otherwise input the correct response and then press the space bar.

In addition, you will also be performing a long term memory task. At random intervals throughout the session you will see the following message in the box at the lower right portion of the screen: "ENTER CURRENT VALUE". Your task will be to mentally subtract three from the previous number and input this resulting value into the computer using the number keys at the right of the keyboard. The experimenter will tell you which number you should start subtracting from at the beginning of the session.

It is important to keep one thing in mind: During the session the decision making task is of primary importance AND SHOULD BE PERFORMED BEFORE THE LONG TERM MEMORY TASK, should they occur together.

Please take this opportunity to go over these instructions with the experimenter and ask any questions you might have.

C4 (Primary Task - Tracking Level of Aiding - Manual)

During the next hour you will be asked to perform two tasks -- a tracking task and a long term memory task. In the tracking task you will be asked to keep the cursor within the center of the solid triangles in the middle of the tracking bar. This is accomplished by using the "<" key when the cursor leaves the center and goes to the right side of the tracking bar and by use of the ">" key when the cursor leaves the center and goes to the left side of the tracking bar.

In addition, you will also be performing a long term memory task. At random intervals throughout the session you will see the following message in the box at the lower right portion of the screen: "ENTER CURRENT VALUE". Your task will be to mentally subtract three from the previous number and input this resulting value into the computer using the number keys at the right of the keyboard. The experimenter will tell you which number you should start subtracting from at the beginning of the session.

It is important to keep one thing in mind: During the session the tracking task is of primary importance AND EVERY EFFORT SHOULD BE MADE TO INSURE THAT THE CURSOR IS WITHIN TOLERANCE BEFORE THE LONG TERM MEMORY TASK IS ATTENDED TO.

Please take this opportunity to go over these instructions with the experimenter and ask any questions you might have.

C5 (Primary Task - Tracking Level of Aiding - Advice)

During the next hour you will be asked to perform two tasks -- a tracking task and a long term memory task. In the tracking task you will be asked to keep the cursor within the center of the solid triangles in the middle of the tracking bar. This is accomplished by using the "<" key when the cursor leaves the center and goes to the right side of the tracking bar and by use of the ">" key when the cursor leaves the center and goes to the left side of the tracking bar.

During this session you will be receiving help in performing the tracking task. Throughout the session you will see a number located at the top of the solid triangle which will tell you how many "units" you are away from the center. For example, if the number at the top of the solid triangle is "-2" this tells you that you are 2 units to the left of center and you need to press the ">" key in order to get back within tolerance. In addition, there will be a small dot immediately below the cursor which will tell you where the cursor will be if you do not input any corrections.

In addition, you will also be performing a long term memory task. At random intervals throughout the session you will see the following message in the box at the lower right portion of the screen: "ENTER CURRENT VALUE". Your task will be to mentally subtract three from the previous number and input this resulting value into the computer using the number keys at the right of the keyboard. The experimenter will tell you which number you should start subtracting from at the beginning of the session.

It is important to keep one thing in mind: During the session the tracking task is of primary importance AND EVERY EFFORT SHOULD BE MADE TO INSURE THAT THE CURSOR IS WITHIN TOLERANCE BEFORE THE LONG TERM MEMORY TASK IS ATTENDED TO.

Please take this opportunity to go over these instructions with the experimenter and ask any questions you might have.

C6 (Primary Task - Tracking Level of Aiding - Total)

During the next hour you will be asked to perform two tasks -- a tracking task and a long term memory task. In the tracking task you will be asked to keep the cursor within the center of the solid triangles in the middle of the tracking bar. In order to help you do this the computer will perform the tracking task, however, you still will be required to monitor the position of the cursor to insure that the computer is performing the task properly (it is possible that it may cause the cursor to go out of tolerance so you must still watch it constantly). In order to insure that you are monitoring the status of the tracking task, you will be required to press the space bar at random intervals throughout the session when the question mark appears directly below the bottom solid triangle.

In addition, you will also be performing a long term memory task. At random intervals throughout the session you will see the following message in the box at the lower right portion of the screen: "ENTER CURRENT VALUE". Your task will be to mentally subtract three from the previous number and input this resulting value into the computer using the number keys at the right of the keyboard. The experimenter will tell you which number you should start subtracting from at the beginning of the session.

It is important to keep one thing in mind: During the session the tracking task is of primary importance AND EVERY EFFORT SHOULD BE MADE TO INSURE THAT THERE ARE NO QUESTION MARKS ON THE SCREEN WHILE THE LONG TERM MEMORY TASK IS ATTENDED TO.

Please take this opportunity to go over these instructions with the experimenter and ask any questions you might have.

Appendix B. Complete Source of Variation Tables
(Secondary Task Data)

Table 1. Univariate Statistics - Dependent Variable: LOGRT

Source	SS	df	MS	F	p
AIDING	2.25611	2	1.12805	4.67	.0135
TASK	2.47765	1	2.47765	10.67	.0023
SESSION	0.95341	2	3.23189	4.65	.0137
TASK*AIDING	0.48906	2	0.24450	1.01	.3704
AIDING*SESSION	0.48960	4	0.12240	0.69	.6018
TASK*SESSION	0.01823	2	0.00910	0.51	.6005
TASK*AIDING*SESSION	0.05970	4	0.01490	0.84	.5034
SUBJ(TASK*AIDING)	13.05190	54			
SUBJ*SESS(TASK*AID)	1.92132	108			

Table 2. Univariate Statistics - Dependent Variable: SD

Source	SS	df	MS	F	p
AIDING	26.21402	2	13.10700	25.19	.0001
TASK	8.85736	1	8.85736	5.71	.0204
SESSION	25.57079	2	12.78540	25.19	.0001
TASK*AIDING	5.32197	2	2.66100	1.71	.1896
AIDING*SESSION	2.50699	4	0.62670	1.23	.3005
TASK*SESSION	0.64246	2	0.32120	0.63	.5331
TASK*AIDING*SESSION	2.69990	4	0.67500	1.12	.3519
SUBJ(TASK*AIDING)	83.79226	54			
SUBJ*SESS(TASK*AID)	54.82426	108			

Table 3. Univariate Statistics - Dependent Variable: LOGREP

Source	SS	df	MS	F	p
AIDING	6.46378	2	3.23190	4.65	.0137
TASK	2.54839	1	2.54839	3.67	.0607
SESSION	4.13042	2	2.06520	4.73	.0107
TASK*AIDING	0.22332	2	0.11170	0.16	.8519
AIDING*SESSION	2.32616	4	0.58150	1.33	.2627
TASK*SESSION	1.15470	2	0.57740	1.32	.2707
TASK*AIDING*SESSION	0.48372	4	0.12090	0.28	.8923
SUBJ(TASK*AIDING)	37.50170	54			
SUBJ*SESS(TASK*AID)	47.14822	108			

Table 4. Univariate Statistics - Dependent Variable: LOGMISS

Source	SS	df	MS	F	p
AIDING	6.69793	2	3.34900	2.50	.0918
TASK	7.31962	1	7.31962	5.46	.0232
SESSION	17.94045	2	8.97020	32.18	.0001
TASK*AIDING	9.09792	2	4.54900	3.39	.0410
AIDING*SESSION	0.50418	4	0.12600	0.45	.7706
TASK*SESSION	0.47118	2	0.23560	0.85	.4323
TASK*AIDING*SESSION	3.01504	4	0.75380	2.70	.0342
SUBJ(TASK*AIDING)	72.43214	54			
SUBJ*SESS(TASK*AID)	30.10423	108			

Table 5. Univariate Statistics - Dependent Variable: LOGINC

Source	SS	df	MS	F	p
AIDING	1.11667	2	0.55830	0.47	.6248
TASK	0.26149	1	0.26149	0.22	.6393
SESSION	2.28306	2	1.14150	7.59	.0008
TASK*AIDING	0.23944	2	0.11970	0.10	.9034
AIDING*SESSION	0.38676	4	0.09670	0.64	.6329
TASK*SESSION	0.25629	2	0.12810	0.85	.4293
TASK*AIDING*SESSION	0.18625	4	0.04660	0.31	.8709
SUBJ(TASK*AIDING)	63.54349	54			
SUBJ*SESS(TASK*AID)	16.23674	108			

Table 6. Univariate Statistics - Dependent Variable: LOGSIG

Source	SS	df	MS	F	p
AIDING	1.88705	2	0.94350	0.35	.7091
TASK	0.23783	1	0.23783	0.09	.7689
SESSION	10.97599	2	5.48800	15.93	.0001
TASK*AIDING	4.42597	2	2.21300	0.81	.4495
AIDING*SESSION	0.94653	4	0.23660	0.69	.3005
TASK*SESSION	1.78495	2	0.89250	2.59	.0796
TASK*AIDING*SESSION	0.93671	4	0.23420	0.68	.6073
SUBJ(TASK*AIDING)	147.2620	54			
SUBJ*SESS(TASK*AID)	37.1945	108			

Micheline Y. Eyraud

Micheline Y. Eyraud was born on 7 December 1950 in Santa Rosa, CA. She is a Lieutenant Commander, U.S. Navy Aerospace Experimental Psychologist. Other academic degrees she has earned are:

M.S. Management	1980	Troy State University
M.A. Psychology	1978	California State University, Fullerton
B.A. Psychology	1974	University of San Francisco

Articles she has published are:

1. Eyraud, M.Y., Bagrash, F.M. & Stoffer, G.R. (1983). Figural aftereffects: An explanation in terms of multiple mechanisms in the human visual system. Pensacola, FL.: Naval Aerospace Medical Research Laboratory Monograph 30.

2. Eyraud, M.Y., Owens, J.M. & Goodman, L.S. (1984). Technical evaluation of the Naval Automated Pilot Aptitude Measurement System (NAPAMS). Pensacola, FL.: Naval Aerospace Medical Research Laboratory Special Report 85-3.

3. Eyraud, M.Y. & Borowsky, M.S. (1985). Age and pilot performance. Aviation, Space and Environmental Medicine, 56, 553-558.

She is a member of the Human Factors Society and the Association of Aviation Psychologists.