NASA Conference Publication 2103

NASA-CP-2103 19790025667

Models of Human Operators in Vision Dependent Tasks

Proceedings of two tutorial seminars
held in Boston, Massachusetts,
October 29 - November 1, 1979

NASA Conference Publication 2103

Models of Human Operators in Vision Dependent Tasks

Marvin C. Waller, Editor

Proceedings of two tutorial seminars
held in conjunction with the Human
Factors Society 1979 Annual Meeting
Sponsored by the HFS Visual Performance
Technical Group and NASA Langley
Research Center, Hampton, Virginia,
and held in Boston, Massachusetts,
October 29 - November 1, 1979

National Aeronautics
and Space Administration

**Scientific and Technical
Information Branch**

1979

FOREWORD

The Visual Performance Technical Group of the Human Factors Society (HFS) was organized in October of 1978 at the Society's Annual Meeting. One of the events that had transpired earlier and that contributed to the thinking of many of those involved was the 1977 HFS tutorial seminar organized by Dr. H. Zwahlen of Ohio University and Dr. Thomas Rockwell of Ohio State University on the subject of operator scanning behavior. Discussions in this tutorial led to the conclusion that insufficient progress was being made in developing a body of theory to describe how operators visually absorb information from displays and explain operator scanning behavior. Many of the participants in that 1977 tutorial are now active members of the Visual Performance Technical Group. The concerns expressed in the 1977 tutorial seminar have, in part, supplied the motivation to organize these tutorial seminars on modelling of the human operator in a vision dependent task.

The focus of the Visual Performance Group in this modelling effort is on a submodel to represent how the visual system acquires information from a display or scene and how this process contributes to and interacts with the overall operation of the system. Since operator information acquisition does interact with the overall control of the system, it is clear that the visual information channel should be studied and modelled within the context of the larger framework of which it is a part. Therefore, visual performance researchers should be cognizant of the structure of existing large-scale, computer-implemented models so that their research contributions can be properly integrated into the model structure.

NASA Langley Research Center, through its Terminal Configured Vehicle (TCV) program, is engaged in research and development to improve airborne systems and operational procedures, with emphasis on terminal area operation. A major area of this research is the proper integration of the pilot as an integral part of the total aircraft system. Numerous simulator-based and flight investigations have been conducted in conjunction with this research program to understand factors in the display and control systems which contribute to pilot performance, workload, and overall system reliability and safety. It is highly desirable in the long term to integrate the results of these investigations into a theory-based model of the total flight system including the pilot. Such a model would benefit the program's future research efforts by allowing many studies to be made on an analytical basis and reduce the requirement for extensive empirical investigations. The model would additionally be a catalog of the wealth of information being acquired through the research efforts of this program.

Interest in the use of computerized models in flight management research is not an all together new idea at NASA Langley Research Center nor in the TCV program. Some applications of the optimum control model in display research have been pursued under contract with Bolt Beranek and Newman, Inc. Also, Langley has sponsored an effort to develop a timeline analysis workload evaluation program under contract with Boeing Company of Seattle. NASA's involvement in these tutorials results from its continuing interest in the use of modelling techniques as a research tool.

Two tutorial seminars have been organized to cover related topics in this field of human factors research. The first session, entitled "Modelling of the Human Operator," is a review of some of the research and development conducted to date to model the complete operator functions during vision dependent tasks. The second tutorial, entitled "Modelling of Visual Information Processing," will address the structure and details of the submodels on visual information processing included in the large computerized models covered in the first tutorial. Some additional approaches to visual information evaluation are also presented.

Recognizing the significant level of effort required to develop such models, dissemination of information on the operation and current status of existing models and their application to the area of visual performance research is the primary goal of these tutorials. The two subject tutorials have been organized to present an overview of work in the human operator and vision information processing modelling areas. Some choices had to be made on what material to present since the time allowed would not permit even a reasonable overview of all of the known work in these areas. It is anticipated, however, that enough details of the modelling efforts of the selected works will be included in the seminar to permit the participants to develop a realistic concept of some of the pragmatic considerations involved in selecting and using a model of the human operator in vision dependent tasks.

CONTENTS

A SURVEY OF APPLIED PSYCHOLOGICAL SERVICES' MODELS OF THE HUMAN OPERATOR

Arthur I. Siegel and J. Jay Wolf
Applied Psychological Services, Inc.

SUMMARY

A historical perspective is presented in terms of the major features and status of two families of computer simulation models developed by Applied Psychological Services in which the human operator plays the primary role. Both task oriented and message oriented models are included.

Two other recent efforts are summarized which deal with visual information processing. They involve not whole model development but a family of subroutines customized to add the human aspects to existing models.

A global diagram of the generalized model development/validation process is presented and related to 15 criteria for model evaluation.

INTRODUCTION

Our goal in this report is to distill the essence of over 21 years of continuous effort at Applied Psychological Services to develop and validate digital computer models. Our primary area of concentration has been the emulation of systems in which the key element is the human who operates and/or maintains equipment systems. Over 30 projects (contracts) have been devoted to the efforts reported and the models have been applied to systems for all of the military services as well as NASA and industry.

The models are summarized in three families as outlined in Table 1. Three "task oriented models" were developed to simulate man-machine systems. Their major distinguishing feature is the size of the crews which they simulate. They all operate on input data describing a list of tasks or subtasks which the operator(s) or groups of operators perform. Each task is simulated sequentially by the logic of the model implemented by a computer program so as to allow operators to: work independently or together, wait for each other, talk to each other, monitor and operate controls and displays, wait for equipment, skip nonessential subtasks if the operators are busy, make decisions which can alter the subtask sequence, recycle if required in the event of an operator failure, become partially or completely incapacitated, and respond to unexpected failures and emergencies.

The major attribute of the two models in the second family is that they track each message in a communications system. Embedded in each of these, however, is a task oriented module--a miniature of the first model family.

Table 1

Model Families

Family	Family Name	Models
I	Task Oriented Models	One-two Man Model Intermediate Crew Model Large Group Model
II	Message Oriented Models	Tactical Operations Message Handling Field Exercise Monitoring
III	Human Oriented Subroutines for Models	Target Detection & Classification Facility Defense Combat

The paper also describes two recent efforts not directed at development of an entire model--because the model in question had already been developed. In these cases, it was realized after model completion that the human aspects were not adequately simulated. As a result, a group of subroutines was developed to add the realism of a human element in tasks such as scan, detection, and target classification.

All of these models were designed for use in simulating difficult or untried missions--those in which the operator's physical and mental limitations may play an important part in the ability of the man-machine system to perform its function. The conceptual design of each member of each model was based on our unwavering belief that human behavior in a dynamic environment cannot be validly represented by deterministic methods. Last, the paper discusses model validation and a set of validation criteria.

FAMILY OF TASK ORIENTED MODELS

Consider first the family of task oriented models--the mainstay in our arsenal of simulation tools.

All models in this family were prepared to simulate men operating and/or maintaining equipment. All have major simulation variables to reflect the realities of the equipment, the mission itself, and one or more important time functions. Yet they all possess, in addition (and this represents their distinctive feature), psychological and social variables pertaining to the operator or to groups of operators. Examples of these are stress, orientation, proficiency, mental load, and fatigue. Flexibility in simulation is provided in the models through the ability to allow parametric variation of such factors as the speed of the operators, their stress breaking points, and mission time limits. In addition to the more common or system oriented results such as equipment reliability, working hours, and operator failures which one has come to expect from computer models, these three models generate data on personnel performance, morale, cohesiveness, goal orientation, and man-machine system efficiency. All yield computer output tabulations reflective of the man-machine system under study in order to predict system "performance," personnel overloads, periods of unusual stress and excessive delays, distributions of how mission time is spent, a variety of end-of-mission conditions, and implications of manning strategies.

The principal features and differences between the three models are shown in Table 2.

The 1-2 man model (the first entry) simulates one or two operators and accommodates up to 300 individual actions by each operator. Each such operator action which would require a few seconds or minutes of operator performance is simulated by the computer in about 3 milliseconds. Consequently, 100 computer iterations of a maximum task (300 actions for each operator) take about two to four minutes of computer time.

In the intermediate model, a crew of up to 20 men may be simulated. It handles the case of multi-day missions in which the times of individual events

Table 2

Major Features of the Task Oriented Models

NAME	Number of Personnel Simulated	Duration of Mission	Duration of Tasks, Events	Number of Tasks, Events
One-Two Man Model	1-2	minutes, hours	seconds, minutes	up to 300
Intermediate Crew	3-20	a few days	tenths of hours	80 per day
Large Group	20-100	many days	tenths of hours, hours	100 per day

are measured in minutes or hours. This is accomplished by processing tasks performed by groups of one or more men. Here, the computer simulates each of these longer events in about 20 milliseconds. In this case, 100 iterations of a maximum mission (80 crew events per day) for, say, a five day mission, take about 10 to 15 minutes of computer time.

In the largest model, a crew of from 20 to as many as 100 men may be simulated. The mission is composed of work units which may be minutes or hours in duration and the total mission may last for several dozen days. The limit here is principally a practical one based on computer running time.

Table 3 itemizes the principal concepts of the task oriented models. The elements in the table are presented principally to indicate the scope of models. Listed under "Major Human Features and Variables" are the types of functions the model can handle, i.e., the principal variables it considers. The operator-oriented variables selected for incorporation into the models are those which psychologists have determined to be influential on the performance of an individual or of a closed social group. Listed under "Principal Parameters" are the items that the systems analyst who is using the model can vary--that is, he can select values for each parameter for each computer run of the model. Under "Major Outputs" the principal categories of printed computer output are shown.

THE ONE-TWO MAN MODEL

This 1-2 man model has been in active use for almost two decades. It has been tested against real life and against laboratory controlled criteria, and has found to give reasonable, interesting, and valuable results.

The major results from using this model are:

- the probability of success--that is, the percentage of times that the prescribed sequence of subtasks was completed within the time limit.
- the shape of the stress function during the simulation.
- the distribution of time spent working, waiting, and in repeating work not properly performed.

This model is still in active use by a variety of government and commercial users. Several applications have been made by the originators and many by others including landing an aircraft on an aircraft carrier, launching an air-to-air missile, an inflight intercept of an enemy aircraft, and simulation of an inflight refueling operation. These represent both one and two operator simulations.

A preprocessor program has been developed which yields another version of the model. It calculates adjustments to data normally provided as model input prior to simulation. In this case, the effect of radiation exposure on performance time and success probability is determined for man and/or machine degradation.

Table 3

Features of the Task Oriented Models

NAME	1-2 Man Model	Intermediate Crew	Large Group
Major Human Features and Variables	proficiency time stress cohesiveness decision making	proficiency time stress & mental load fatigue & sleep physical capability intelligence learning aspiration supervisors expectation sickness/incapacity	proficiency time stress fatigue norms & goals cohesiveness social pressure learning morale
Principal Equipment/Environment Variables & Features	reliability equipment response	reliability (failures) emergencies communications task postponement consumables hazard level	reliability emergencies communications task postponement consumables mean time to repair
Major Outputs	success probability stress profile work, idle, failure time performance task repetition time tasks failed, ignored	success probability stress profile work, repair, idle, failure time MTBF, MTTR, availability performance adequacy tasks failed, ignored	success probability crew efficiency morale tasks failed, ignored work, idle, failure, repair time
Principal Parameters	operator time limits operator stress threshold operator individuality factor nuclear radiation dose	 operator stress threshold crew work pace work day length acceptable performance level crew qualification level	crew size, increment operator proficiency operator stress threshold acceptable performance level work hours per day crew composition

Table 3 (con't)

Application	aircraft landing	USCG Patrol Gunboat	FBM submarine stationary underwater station
	inflight intercept sonar inflight refueling missile launching	AN/SQS-76 Sonar	nuclear missile
Current Use	yes	yes	no

The 1-2 man model was also adapted for subject-to-computer dynamic, on-line interaction. Here, the model simulates the performance of perceptual-motor acts and routine operations while one or two subjects, who are seated at independent graphic video display terminals, perform selected task elements.

THE INTERMEDIATE CREW SIZE MODEL

In the intermediate crew size model, a crew of up to 20 men and multi-day missions may be simulated.

The model, which is heavily group oriented, includes the use of several types of statistical distributions. For example, numbers are drawn from an exponential distribution to determine the time(s) that equipment failures are to be imposed, from a rectangular distribution to determine placement of the emergencies in the list of events each day, from a normal distribution for estimates of mean performance time, and from a poisson distribution to select the number of days duration of sickness of operators. The general simulation sequence is:

- Crew formation--The model identifies each crew member and assigns a value for speed, aspiration, and competence.
- Daily schedule generation--This is done by interspersing prearranged mission events with unforeseen repairs and emergencies.
- Personnel assignment for each event sequentially--Here, the model selects an individual man or a group of men to accomplish the work of each event, ignoring events depending on the essentiality of the events and other factors. The leader of the group is also assigned.

- Event simulation--Calculation of conditions existing prior to the event, and how well and how quickly the assigned men accomplish the work event, and selection of the next course of action.

- Update--Modification of the numerical status of psychosocial and other variables as a result of group performance.

- Output--Selection and printing the value of key variables and summarizing end conditions for each event, each day, each mission iteration, and a summary of all iterations.

The intermediate model was tested to assess its sensitivity and to estimate its validity--that is, the extent to which the model's output agrees with independent criterion data. The mission selected for simulation was that of an 82-foot U.S. Coast Guard boat responsible for patrol of Vietnamese waters. This four-day mission involved a heavy schedule of investigating various river craft, boarding a suspected boat for search operations, navigation, steering, engine monitoring, cleanup, clerical work, preventive maintenance, administrative duties, meal preparation and eating--60 to 65 events in all.

This model has been in almost continuous use since its initial development and validation in 1969. It has been improved by developing a version which simulates equipment, human, and system reliability oriented calculations. As a result, the model yields a number of output numerics believed to possess considerable relevance to human and system availability and reliability prediction. These include: human reliability, availability, and MTTR; equipment reliability, availability, MTTR, and MTBF; and system reliability, availability, and MTTR. This use advances the role of the reliability engineer from that of an actuarial to that of a true system designer or system design advisor who provides an active and ongoing contribution to the total system design and effectiveness assurance process.

This model was also the subject of a set of parametric computer runs for the purpose of developing a set of human tradeoff curves. These were published to show, in handbook format for design engineering use, the relative impact of some human oriented variables on system performance.

THE LARGE GROUP MODEL

In the large group model, the mission is composed of work units, each of which may be minutes or hours in duration, and the total mission may last for several dozen days. Since this model is concerned with group performance, the inputs to the model are principally concerned with group oriented variables salient to behavior. In this mode, variables such as crew morale, cohesiveness, operator orientation, proficiency, performance time, overtime, communications, sickness, and system effectiveness are computed.

In the use of this model, we conceive of supervisors and workers who together form a relatively large crew. In the performance of each job, the computer "selects" the proper number of appropriately skilled men to form a group who "accomplishes" the work in a time and under other conditions which are numerically calculated.

The large group model is the only one of the family in which the computer is programmed to calculate the crew size. This is an optional feature so that simulation runs can be made with the crew composition prespecified, or if left unspecified, simulation will be initiated with what is considered a minimum crew as determined by the logic of the model. Then additional simulations are performed successively--each time with a larger crew. For each increase, the computer selects the most needed man or men to be added to the crew. This process continues uninterrupted until a preset parametric limit on crew size is exceeded, or until the crew reaches a size which eliminates the need for overtime work.

Sensitivity runs made on this model were based on application to a fleet ballistic missile submarine. A series of 10 day missions was simulated with crew sizes which ranged from 33 to 44 men working at five stations, using actual data available during the FBM planning stage. The model was run through cruise operations, stationary submerged operations, and emergency drills which are representative of typical missions. The results compared favorably with actual system mission data. In particular, predictions from the model of system effectiveness, in the composition of the crew, and in its proficiency agreed very well with quantitative data as well as qualitative opinion summarized from interviews with officers of FBM submarines.

Operational validation of this model was completed using data from underseas craft of the 627 class of submarines. Numerous computer simulations of a 21-day mission were made with crew sizes varying from 48 to 61.

It is noted that this model requires extensive data input, and possibly, as a result, this is the only model of the three task oriented models which has had no recent activity.

MESSAGE ORIENTED MODELS

Two models were developed to simulate those aspects of systems whose primary purpose is the operational handling of messages. These models keep track of each message text processed in the system and also simulate the acts and behaviors of operations personnel as they receive, prioritize, code, and enter messages in the system. The models are completely general and allow for the simulation of personnel of different competencies and stress tolerances, along with a variation in message load and content.

These models combine the effects of such features as message generation and queuing, detailed message processing procedure, error rates, and personnel characteristics, along with stochastic variations to yield predictions of system performance. As in the task oriented family, the basic nature of both models is stochastic. As a result, a number of repetitions is required to produce a stable result.

Along with the simulation of human message processing, the models include the simulation of the computer embedded in each of the target systems. Some global information about both of the message processing models is given in Table 4. Both models handle multiple message types of varying priorities.

The first of the models, initiated in 1972, was directed to the simulation of message processing within the Tactical Operations System. TOS is an automated, secure information processing system designed to assist military commanders and their staffs at Field Army, Corps, and Division levels in the conduct of tactical operations.

There are up to four sequences of task elements provided to represent the tasks executed by an operator in performing his duties. Each sequence has the capacity of up to 20 task elements. The model handles up to 6 men of 2 types, 4 types of operator errors, 7 types of messages, 4 message priority classifications, and a shift length of up to 12 hours.

At the start of simulation for a new TOS shift, a backlog subroutine generates data representing messages in the action officer's "in-box" at the start of each iteration. A message generation subroutine develops data representing messages which will arrive during the coming hour. These are merged with the backlog in order by time of arrival, and each message of this hourly message queue is processed in turn by a single selected operator. The operator stress and aspiration conditions applicable to that situation are calculated next. The detailed task element-by-task element simulation for the message and operator selected is accomplished by a subroutine which manipulates mission task analysis data in a way very similar to that used in the 1-2 man model described earlier.

Output from the model includes detailed and summary tabulations including an hourly summary, shift summary, and run summary.

The simulation run summary includes sections for manpower utilization, message processing time, overall efficiency indicator, workload summary, and error summary. In this form, the original sensitivity tests were run, and the model was validated against a set of error data collected from an independent source. A high degree of correspondence with the independent data was found.

In a follow-on effort, the model was modified to operate in an interactive time sharing mode, allowing the experimenter and one or more subjects to interact in a "conversational" mode with the model and to enter data "on line." Various extensions of the original model were also made at this time. A variant of the original model was also included which allowed collection of data during an experiment in which one or more actual operators performed a part of the process and the computer simulated the remainder of the TOS activity.

More recently, the TOS model was adapted for the UNIVAC 1108 computer, and several new capabilities were added which increase the realism of the simulation. It was modified to exchange data with two other independent Army computer models

Table 4

Description of Message Handling Models

System Simulated	Tactical Operational System	Military Exercise Control/ Evaluation System
Program Name	MANMOD	NETMAN
Maximum Number of Men Simulated	6	57
Types of Personnel	2	4
Major Input Parameters	Shift Length Number of Personnel Error Rates Operator Characteristics Speed Precision Aspiration Stress Message Characteristics	Shift Length Number of Personnel Error Rates Operator Characteristics Speed Precision Aspiration Stress Message Characteristics Network Data
Major Output	System Effectiveness Time Worked Operator Stress, Aspiration Message Processing Statistics Errors	System Effectiveness Time Worked Operator Stress, Aspiration Message Processing Statistics

in such a way as to maximize the strong points of each of the models.

The end result is the ability to answer questions such as:

- How does system effectiveness vary as a function of message load, operator level of aspiration, message arrival time distribution, or personnel proficiency?
- What is the effect of increasing or decreasing the manning level or personnel proficiency?
- How much stress was on the operators during the performance of the work of each hour?
- What is the error rate for various message types and for various mannings and personnel attributes within manning?
- How much time was spent, on the average, processing each type of message?

The Army Field Exercise Model

Most of the techniques used in the TOS model were utilized in developing an expanded model for simulating the message handling aspects of Army field exercises in which up to 27 referees, 27 radio operators, and 3 controllers interact in a fixed, closed loop network of communication lines while sharing time on a central computer. Messages introduced into the system are prepared, processed, and entered into the computer by various personnel and delivered to the controllers for evaluation.

Each computer run of the model represents a simulation of up to 10 hours in duration, in which up to 2000 messages can be processed. In this model, each operator type has its own task analysis.

This model has recently been the subject of both sensitivity and validation testing. A series of 59 computer runs enabled statistical test on the effects of a variety of personnel and workload variables, manpower configurations, and task variables. The results were found to be reasonable and appropriate; the most influential variables were operator speed, operator precision, and network configuration. The psychological factors (stress, aspiration level) exerted a much less powerful effect on output.

HUMAN ORIENTED SUBROUTINES

Two other developments have recently been completed which led to the specification of several computer subroutines designed to provide the capability to simulate the personal, psychosocial, and group interactive aspects involved in the target system. The subroutines are designed to be suitable for interfacing with a parent program which simulates other aspects of the system.

The first effort produced four types of different, yet related, computer subroutines or modules. Each of these was conceived to operate as a part of a global computer program whose goal is to simulate the principal ground based, man-machine operations involved in the AN/UPD-X system. In this USAF system, video type displays present replicas in real time from processed data sensed by a side looking radar, mounted in a USAF reconnaissance aircraft.

The subroutines, defined here are those human oriented functions, involved with the capability of simulated operators to perform basic tasks:

- In the SCAN/DETECT Module the operator scans a cathode ray tube (CRT) screen for the presence of targets and detects targets.
- The CLASSIFY Module involves determining which type of target has been detected.
- The DECISION Module simulates operator decision making.
- The COMMUNICATIONS Module involves simulation of inter-operator communications during AN/UPD-X operations.

Each of these subroutines determines the amount of operator time required in the simulated performance of these tasks and determines whether or not the simulated operator(s) performed these tasks adequately (i.e., sucessfully or unsuccessfully).

The AN/UPD-X system was in the design or "evaluation of alternatives" phase during the model development period. As a result, the human oriented subroutines were developed in a sufficiently general way to allow their use during comparative simulation of alternative AN/UPD-X system designs--even those developed by different industrial contractor teams including different AN/UPD-X equipment configurations and diverse operator sequences. Generality was a goal in these module designs--so that the modules will be valid across various equipment and AN/UPD-X system designs developed by several USAF contractors. This objective was achieved in that a user of these modules need only modify inputs to subroutines in order to accommodate system oriented feature differences such as:

- radar coverage area
- CRT display characteristics and size
- target types
- operator ability
- mission time
- communications load

NUCLEAR FACILITY ATTACK SIMULATION

The other effort leading to human effect modules was directed to a model which pitted an attacking force against a force defending a nuclear facility. The hostile intruder attack had been simulated by a hostile attack simulation model which previously had no human behavioral features.

Four features were selected because of their important effect on human performance and were incorporated:

- effects of nuclear radiation
- visual effects of illumination (light level)
- effects of stress
- group cohesiveness effect

MODEL VALIDATION CONCEPTS

Emshoff and Sisson (1970) in a discussion of model validity concluded that: "the only possible evidence of validity for a simulation model that has been developed specifically for a situation is that the model has made satisfactory predictions in the past." They suggested five "preliminary criteria for evaluating first time models" as described by Hermann (1967). These five are identified by an asterisk in the more comprehensive list of 15 criteria for evaluating a simulation model which are displayed in Table 5. These criteria are not necessarily mutually exclusive. Some are overlapping, but all are considered important in some sense and/or for some classes of models.

In order to place these criteria into some perspective and to view the sequential steps through which our models pass, consider Figure 1, which attempts to tie together the various model development/validation phases with these 15 criteria for model evaluation. This figure displays the major steps (large rectangles) from concept and model requirements derivation through the situation in which the model can be considered for decision aiding and eventually for decision making. The 15 numbered vertical arrows, representing the 15 criteria, show that each step in the process yields some measure of utility, feasibility, cost, reasonableness, or validity. It is suggested that a model whose design meets the criteria emanating from the model design box be said to be "suitable" (see lowest oval). A model which is programmed and debugged enters a state here called "testable." After sensitivity testing (and the implementation of corrections to the model as required), the model is said to be "reasonable." Following adequate validation testing, the model is termed "valid" or "useable" for decision aiding and, after the experience of use, the model is "operative," "proven," or "effective." The various types of data and information required as inputs to each phase are shown entering from the left with the resulting documentation outputs exiting to the right.

Table 5

Criteria for Evaluating the Utility of a Computer Model

	Criterion	Definition
1.	Internal consistency	Extent to which the constructs of the model are marked by coherence and similarity of treatment
2.	Indifference to trivial aggregation	Potential of the model to avoid major changes in output when input groupings or conditions undergo insignificant fluctuations
3.	Correct prediction in the extreme (predictive or empirical validity)	Extent of agreement (correctness of predictions) between model and actual performance at very high/low values of conditions
4.	Correct prediction in midrange (predictive or empirical validity)	Like above for middle ranges values of conditions
5.	Construct validity	Theoretic adequacy of the model constructs
6.	Content (variable parameter) validity (Fidelity)*	Extent to which the model's variables/parameters match real life conditions
7.	Realism or "face validity"*	Extent to which selected content matches each attribute modeled
8.	Richness of output	Number and type of output variables and forms of presentation
9.	Ease of use	Extent to which an analyst can readily prepare data for, apply, and extract understandable results from the model
10.	Cost of development	Value of effort to conceive, develop, test, document, and support
11.	Transportability-generality	Extent to which different systems, missions, and configurations can be simulated
12.	Cost of use	Value of all effort involving use of model including data gathering, input, data processing, and analysis of results

13.	Internal validity*	Extent to which outputs are repeatable when inputs are unchanged
14.	Event or time series validity*	Extent to which simulation predicts event and event patterns
15.	Hypothesis validity*	Extent to which model relationships correspond to similar relationships in the observable universe

* Approaches to validation defined by Hermann (1967)

FOOTNOTE

The task oriented models were originally developed under contract with the Engineering Psychology Programs and Organizational Psychology Programs, Office of Naval Research. Enhancements for radiation (and other decrement effects) and reliability/availability effects were sponsored by the Aeromedical Research Laboratory, Wright-Patterson Air Force Base and the Naval Sea Systems Command, respectively. The message oriented models were sponsored by the U.S. Army Research Institute for the Behavioral and Social Sciences.

The modules relating to hostile attack on nuclear facilities were developed for Sandia Laboratories, and those relating to electronic processed imagery systems were sponsored by Aerospace Medical Research Laboratory and the University of Dayton Research Institute.

References

Emshoff, J.R. & Sisson, R.L. Design and use of computer simulation models. London: MacMillan, 1970.

Hermann, C. Validation problems in games and simulations. Behavioral Science, 1967, 13, 216-230.

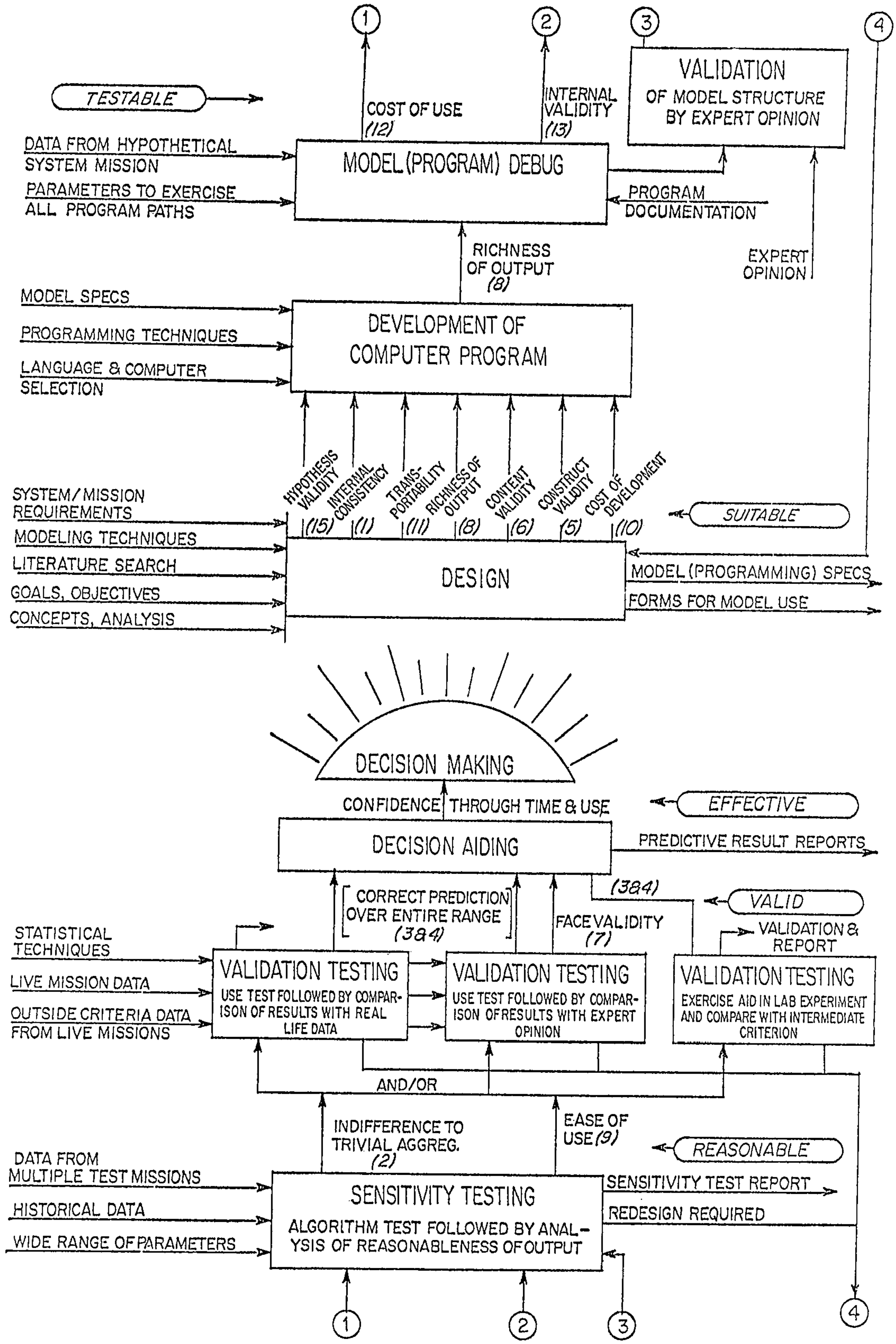


Figure 1.- Steps in model development.

VISUAL PERFORMANCE MODELING IN THE HUMAN OPERATOR SIMULATOR

Melvin I. Strieb
Analytics
Willow Grove, Pa.

ABSTRACT

A brief description of the history of the development of the Human Operator Simulator (HOS) model is presented. Features of the HOS micromodels that impact on the obtainment of visual performance data are discussed along with preliminary details on a HOS pilot model designed to predict the results of visual performance workload data obtained through oculometer studies on pilots in real and simulated approaches and landings.

INTRODUCTION

The HOS model has been under development for approximately 10 years. The concept behind the model was formulated by Wherry (Ref. 1) in 1969. Analytics began the task for formalizing Wherry's ideas and converting them into a functioning model (Ref. 2) in 1971. Development of the basic model was completed in 1976 (Refs. 3, 4, 5, 6, 7, 8) when the model was first applied to a major Naval weapons system (Ref. 9). Since that time, the model has been applied to several other Naval systems (Refs. 10, 11, 12, 13). Each application has resulted in an increasing confidence in the validity and generality of the model and in an expansion of its range of applicability to more and more complex situations.

HOS developed as a result of Wherry's work in the field of crewstation design, test and evaluation. He recognized that the task analyses that were being prepared for the Navy suffered from several major flaws. First, the analyses never adequately expressed what was expected of the operator. Tasks were specified at varying and usually macroscopic levels of detail (e.g., "Pilot acquires and locks on target") and the times assigned to activities were, at best, educated guesses. The analyses would never indicate that the operator was too busy to perform all the assigned functions (though in actual operational situations, the operator might have been) because the analyses were being prepared by equipment manufacturers who had vested interests in making their systems look good. The analyses did not realistically represent either the dynamics of interactions between mission functions or the interactions between the external world and operator activities.

Wherry concluded that, since there were not standards that the Navy could apply to ensure an unbiased and consistent evaluation, the structure of a task analysis, its level of detail, and its insensitivity to variations in crewstation design did not permit the realistic evaluation of design alternatives. Proposed designs could still only be evaluated by building mockups, simulators, and prototypes and running subjects through test scenarios. But such studies, in addition to being costly, time-consuming and confounded by inter-subject variability, could only be performed so late in the design process that the results of the studies could have only minimal impact on the ultimate system design.

Wherry proposed the development of a computer simulation model that would be capable of simulating an operator in a complex crewstation to the level of detail needed for realistic evaluation of alternative designs within the context of simulated missions. The model would be capable of producing the types of data that had been obtainable only from man-in-the-loop experimentation. The characteristics of the crewstation, the performance requirements of the operator and the details of the missions that could be specified to the model were to be completely general. The simulation would become a specific operator with specific tasks to accomplish when the analyst supplied a description of the crewstation, the procedures the operator was to follow to utilize the equipment and a description of the behavior of the external world, just as a human being becomes the operator of a system when placed in a crewstation, told how to use the equipment and given a specific job to do.

To facilitate the process of defining the crewstation, the operator procedures, and the external world, an English/FORTRAN-like language -- the Human Operator Procedures (HOPROC) language -- was developed. HOS translates HOPROC statements describing macro-level operator actions into micro-level activities whose performance times are dependent on basic human performance characteristics and the mission dynamics (Ref. 14).

The HOS approach differs significantly from the approaches used in models like SAINT (Refs. 15, 16, 17, 18), Siegel-Wolf (Refs. 19, 20, 21), TLA (Refs. 22, 23) and the various control theory models (Refs. 24, 25). The essence of HOS is an *explicit model of the operator* and of how the operator translates procedural statements into activities. Underlying the HOS model is the assumption that human performance (in general) and the performance of a well-trained operator (in particular) is *explainable* as the concatenation of micro-activities. The performance time for each micro-activity is *predictable* and expressed functionally by the micro-model for that micro-activity. Since the human performance micro-models are based on experimental data, HOS is not only a means of evaluating complex systems, but also a structure within which experimental models of human performance can be tested and evaluated.

THE HOS OPERATOR MODELS

There are five major micro-models in HOS -- an anatomy movement model, an information absorption model, a mental computation model, a decision-making model, and a control manipulation model. These models were developed from analyses of both published and unpublished data on human performance.

Where no data or models were found to exist, "common-sense" models were developed. These models can be modified either as new data becomes available or as specific applications indicate the need for model improvements. The models and the sources from which they were derived are discussed in detail elsewhere (Ref. 14). However, for the purposes of understanding visual performance as modeled in HOS, it is necessary to briefly review the eye movement features of the anatomy movement micro-model and the information absorption micro-model and the HOS models of operator variability and error.

Eye Movement

When a HOPROC instruction (e.g., READ THE ALTIMETER) requires the operator to move his eyes to a specific device, the eye movement micro-model is accessed. This model computes a movement time based on location of the current eye fixation point and the new fixation point. The equations used in this micro-model are based on published experimental data on lateral eye movements (Ref. 26) and data from an unpublished experiment by Wherry and Bittner involving both lateral and convergence movements. Figure 1 indicates the range of eye movement times for situations involving only lateral movements between two fixation points on a plane 71 cm (28 in.) from the operator.

Information Absorption

The HOS information absorption micro-model is dependent on a *hab strength* parameter, derived from Hull's learning theory habit strength concept. Information is absorbed in discrete chunks (*micro-absorptions*). Each micro-absorption increases the operator's confidence (hab strength) until the operator is sufficiently confident in his knowledge of the value of the device, at which time the absorption process is terminated.

Each micro-absorption results in the addition of a *micro-absorption time charge* whose value is dependent on input quantities supplied by the analyst in combination with characteristics of the device (e.g., whether the device is discrete or continuous, how many settings it has, etc.). Figure 2 indicates how hab strength varies as a function of time for four different devices.

Operator Performance Variability

The HOS model views operator performance variability as the result of differences in the performance capabilities of different subjects coupled with differences in operator strategies. Differences in performance capabilities are represented by parametric differences in the functional relationships in the micro-models. Differences in operator strategies are representable as either different decision rules in the operator procedures or as differing prioritizations of the operator procedures. By parametrically varying these quantities, HOS can be used to evaluate both the operator performance required by a system and alternative operator strategies and prioritization schemes. The first type of evaluation (operator performance capabilities) can be useful in the process of screening candidate operators. The latter evaluation (strategies and prioritization schemes) can help to develop training procedures that will ensure that operators are trained to optimally utilize the system's capabilities. Although both of these possible uses of HOS have yet to be explored, they were anticipated in Wherry's original conceptualization of HOS. The former was implied by the "o-state" (operator state) concept that allows variations in the operator performance equations throughout the mission; the latter in the criticality values assigned to different operator procedures that can be (and are) dynamically modified throughout the simulation[1] and in the English-like syntax of the HOPROC language that enables the HOS procedures to be used directly as training materials.

Operator Error

One of the most controversial issues associated with HOS is its model of operator error. To understand this model, it is important to remember that the primary objective for which HOS was developed was the evaluation of the nominal performance of a system by a *well-trained, average* operator. By definition, a well-trained operator is one who carries out instructions "by the book," without omitting a step, making an incorrect decision (based on the decision rules specified in the instruction set), or incorrectly carrying out an instruction. However, this definition does not preclude all sources of operator error. For HOS, the significant sources of operator error are:

(1) Requiring the operator to perform more activities in a given period of time than possible (because of human and/or equipment limitations), thereby causing the operator to "fall behind" in the mission.

(2) Giving the operator an incorrect set of decision rules and/or operating instructions, thereby causing tactical and/or operational errors.

(3) Giving the operator poor displays and/or controls that do not permit information to be read or controlled with sufficient accuracy to permit proper operation of the system, causing errors to occur in carrying out subsequent (or concurrent) operations and/or requiring the operator to invest more time, once again causing the operator to fall behind in the mission.

These types of errors *result* in operator performance errors, but are really failures in the design of the system -- flaws which the human factors engineer must address in proposing design modifications. They are problems created when system designers fail to take into account human performance limitations. Clearly, they are not errors of the same sort as when an operator inadvertently pushes a wrong button -- such errors are either random and of low frequency (in which case it is unfair to use them to evaluate the nominal performance of the system) or caused by working the operator beyond capacity. They are, however, the types of errors that must be engineered out of the system.

VALIDATION

Validation of any complex model (and particularly a Monte Carlo simulation model like HOS) is fraught with difficulties. One can argue that such models can *never* be fully validated -- the best one can hope for is that in specific situations, given well-defined sets of inputs, the model can be shown to produce the outputs that match expectations, experience and available data. The problem is even more complex with a model like HOS because, unlike simulation systems that manipulate the user's model of a situation (i.e., the inputs) according to incontrovertible mathematical formulae, in HOS there is both the HOS model of the operator *and* the user's model of how the system functions and how the operator will utilize it. Both models must be valid for the results of any particular simulation to be valid. But since human behavior is so complex, one can never be sure that all possible circumstances have been fully described and all possible alternatives foreseen. It is therefore almost impossible to validate any specific model.

Notwithstanding these difficulties, efforts have been made to ensure both the validity of the HOS operator model and the reasonableness of the outputs obtained from specific user models. Tests of the validity of the HOS model have involved simulations of specific experiments drawn from the human factors and experimental psychological literature (Refs. 8, 10, 11). User model validations have included simulations of specific Navy crewstations (Refs. 9, 12, 13). Both types of simulations have confirmed the general validity of HOS.

Although comparing model results with experimental data has generally been straightforward, validation of the model in complex military situations has been problematical because of the difficulties associated with attempting to capture all the potentially significant variables in the simulation. The converse of this problem is also true -- one can establish a scenario that can be run through HOS, but it is difficult (if not impossible) to set up real-world situations (e.g., at-sea exercises) that will conform to the hypothetical situations modeled in the simulations. Further confirmation of the HOS model is expected as the result of a series of HOS simulations coupled with laboratory experiments that are currently in the planning stages. These simulations will attempt to ensure the validity of the model (and will determine the values of certain input data quantities needed by the model) for a range of situations of varying complexity commonly experienced in Naval weapons systems. In addition, an effort is currently underway with NASA Langley that will test a HOS pilot model through its conformance with visual performance data collected by Spady and Kurbjun (Ref. 27). Preliminary details on this model are presented below.

An operator can be modeled as timesharing his attention among a set of monitoring procedures designed to keep specific displayed items of information at their nominal values. For example, in the approach phase of an IFR landing, a pilot timeshares his attention among at least eight different instruments simultaneously -- the ADF, the radar altimeter, the horizontal and vertical situation indicators, the barometric altimeter, the airspeed indicator, the clock, and the flight director. HOS enables the analyst to describe the pilot's monitoring behavior by a set of *monitor* procedures. Each instrument has its own monitor procedure, e.g.:

```
DEFINE THE PROCEDURE TO MONITOR THE ALTIMETER.
  IF THE ESTIMATED VALUE OF THE ALTIMETER IS WITHIN LIMITS
  THEN WAIT.[2]
      .
      .
      .
```

Such procedures define the actions that the operator is to perform in order to keep the specified instrument within a predefined (and dynamically modifiable) set of limits. These limits, which are defined around a *desired value* (also dynamically modifiable) can be set to a value of zero, in which case the pilot will act like an optimal controller by continuously taking actions to minimize the error. Alternatively, the limits can be set to some non-zero value, in which case the pilot will only take corrective action when the displayed item exceeds its allowable range of variability.

Monitor procedures are executed periodically with a frequency dependent upon a set of decision rules that are part of the HOS decision-making micromodel. These rules use values of how long it has been since the procedure was last executed, how close the device being monitored is to its desired value and the criticality of the device to determine which procedure to work on next. Thus, if all devices are of equal criticality and at their nominal values, each monitor procedure would be executed once before any procedure was executed a second time. By assigning appropriate criticalities to the devices (or to the monitor procedures, themselves), the analyst can control the frequency with which the procedures are executed. When the value of the device differs from the nominal, the HOS decision-making algorithms will perturb the a priori criticalities (and hence the nominal monitoring frequencies) by an amount dependent on the deviation of each device from its nominal value. These changes in the monitoring frequencies correspond to the effects that one sees in a pilot's performance when certain devices become more critical during certain mission phases or when the pilot dedicates more time to maintaining control over certain items because they are harder to control.

Spady and Kurbjun collected (Ref. 27) oculometer data on pilot eye movements during both actual and simulated approaches and landings. Their data functionally describes the variation in the pilot's perceived criticality under varying circumstances. The data on coupled (i.e., autopilot engaged) approaches, for example, (Figure 3), is indicative of operator monitoring frequencies when the operator has a minimum number of functions to perform, i.e., when all devices remain within their limits and no corrective actions are required by the operator.[3] Their data for uncoupled (autopilot disengaged) approaches (Figure 4) indicates how these frequencies change when additional pilot control functions are added. In HOS, this corresponds to increasing the pilot's hab strength thresholds when the pilot is performing the control functions and to the addition of the control activities defined by succeeding statements in the monitor procedures.

It is expected that the HOS micro-models will produce eye movement data directly comparable to the data obtained by Spady and Kurbjun (Figures 3 through 5).

SUMMARY AND IMPLEMENTATION CONSIDERATIONS

This paper has discussed those aspects of the HOS model pertaining to the modeling of visual performance data and the efforts that are currently underway to confirm the validity of those models.

CDR Norman Lane, Naval Air Development Center, Warminster, Pa., directs the Navy's HOS modeling efforts. The Navy is anxious to encourage others to use the model and will provide access to the model for those wishing to.

HOS consists of three major programs which are in FORTRAN, but use some CDC-specific features. The programs would therefore require some (relatively minor) conversion before they could be used on another computer system. The program is large (it can use $200K_8$ words or more of storage for complex simulations[4]) and, for complex problems, can be expensive to run. However, it offers the potential for substantial savings when used as a substitute for real-time simulations and as a means for obtaining types of data that might be virtually impossible to obtain by any other means. HOS should also be considered as an integral part of the system design process, enabling the human factors engineer to propose, test, and either justify or reject proposed system designs based upon a clear and consistent model of human performance.

FOOTNOTES

[1]Criticalities can be explicitly modified by procedural statements and are implicitly modified by the model's decision-making micro-model.

[2]This statement can also be written as either

IF THE ALTIMETER IS WITHIN LIMITS THEN WAIT.

or

IF ALTIMETER IS OK THEN WAIT.

or in any one of a number of other semantically equivalent forms. The HOPROC syntactical analyzer program translates them all into a standard form for use by the simulator.

[3]These data are only *indicative* of the monitoring frequencies because the Piedmont 737's flown were not equipped with an auto throttle; therefore, the pilot was required to control the airspeed with the throttle.

[4]A version of HOS that uses the CDC Extended Core Storage facility is also available.

REFERENCES

1. Wherry, R.J.: *The Development of Sophisticated Models of Man-Machine System Performance* in Levy, G.W. (ed) Symposium on Applied Models of Man-Machine Systems Performance, Report NR 69H-591, North American Aviation/Columbus, 3 November 1969.

2. Wherry, R.J.: *The Human Operator Simulator,* Naval Air Development Center, Warminster, Pa., 25 May 1971.

3. Strieb, M.I., et al: *Introduction and Overview,* The Human Operator Simulator, Vol. 1. Technical Report 1117-I, Analytics, Willow Grove, Pa., 1965.

4. Strieb, M.I., et al: *HOS Users' Guide,* The Human Operator Simulator, Vol 2. Technical Report 1181-A, Analytics, Willow Grove, Pa., 1976.

5. Strieb, M.I., et al: *HAL Programmers' Guide,* The Human Operator Simulator, Vol. 3. Technical Report 1181-C, Analytics, Willow Grove, Pa., 1976.

6. Strieb, M.I., et al: *HOS Programmers' Guide,* The Human Operator Simulator, Vol. 4. Technical Report 1117-B, Analytics, Willow Grove, Pa., 1974.

7. Strieb, M.I., et al: *HODAC Programmers' Guide,* The Human Operator Simulator, Vol. 5. Technical Report 1117-A, Analytics, Willow Grove, Pa., 1974.

8. Strieb, M.I., et al: *HOS Simulation Descriptions,* The Human Operator Simulator, Vol. 6. Technical Report 1181-B, Analytics, Willow Grove, Pa., 1975.

9. Strieb, M.I., et al: *LAMPS Air Tactical Officer Simulation,* The Human Operator Simulator, Vol. 7. Technical Report 1200, Analytics, Willow Grove, Pa., 1976.

10. Glenn, F.A., et al: *Applications to Assessment of Operator Loading,* The Human Operator Simulator, Vol. 8. Technical Report 1233-A, Analytics, Willow Grove, Pa., 1977.

11. Lane, N., et al: *The Human Operator Simulator: Workload Estimation Using a Simulated Secondary Task,* NATO/AGARD Conference Proceeding No. 216, Aerospace Medical Specialists Panel, Methods to Assess Workload, Cologne, April 1977.

12. Strieb, M.I., et al: *P-3C SS-3 Operator Station Cost Effectiveness Evaluation,* The Human Operator Simulator, Vol. 10. Technical Report 1289, Analytics, Willow Grove, Pa., 1978.

13. Strieb, M.I. and Harris, R.: *Sensitivity Analyses for Cost Effectiveness of P-3C SS-3 Operator Station,* The Human Operator Simulator, Vol. 10 - Part 2. Technical Report 1330, Analytics, Willow Grove, Pa., 1978.

14. Strieb, M.I., et al: *HOS Study Guide,* The Human Operator Simulator, Vol. 9. Technical Report 1320, Analytics, Willow Grove, Pa., 1978.

15. Wortman, D.B., et al: *Simulation Using SAINT: A User-Oriented Instruction Manual,* AMRL-TR-77-61, Aerospace Medical Research Laboratory, WPAFB, OH, 1978.

16. Wortman, D.B., et al: *The SAINT Users' Manual,* AMRL-TR-77-62, Aerospace Medical Research Laboratory, WPAFB, OH, 1978.

17. Duket, S.D., et al: *Documentation for the SAINT Simulation Program,* AMRL-TR-77-63, Aerospace Medical Research Laboratory, WPAFB, OH, 1978.

18. Duket, S.D., et al: *Analyzing SAINT Output Using SPSS,* AMRL-TR-77-64, Aerospace Medical Research Laboratory, WPAFB, OH, 1978.

19. Siegel, A.I. and Wolf, J.J.: *A Technique for Evaluating Man-Machine Systems Designs,* Human Factors 3, 1961.

20. Siegel, A.I. and Wolf, J.J.: *Man-Machine Simulation Models,* John Wiley and Sons, New York, 1969.

21. Siegel, A.I., et al: *Techniques for Evaluating Operator Loading in Man-Machine Systems: Evaluation of a One or Two Operator Evaluative Model Through a Controlled Laboratory Test,* Applied Psychological Services, Wayne, Pa., 1962.

22. Miller, K.H.: *Timeline Analysis Program (TLA-1): Final Report,* NASA CR-144942, NASA-Langley Research Center, Hampton, VA, 1976.

23. Anderson, A.F. and Miller, K.H.: *Timeline Analysis Program (TLA-1): User's Guide,* Report D6-42377-3, Boeing Commercial Airplane Company, Seattle, WA, 1976.

24. McRuer, D.T. and Krendel, E.S.: *Mathematical Models of Human Pilot Behavior,* AGARD Report AGARD-AG-188, 1974.

25. Pew, R.W., et al: *Critical Review and Analysis of Performance Models Applicable to Man-Machine Systems Evaluation,* Technical Report 3446, Bolt, Beranek, and Newman, Cambridge, MA. 1977.

26. Dodge, R. and Cline, T.S.: *The Angle Velocity of Eye Movements,* Psychology Review Vol. 8, p. 145-156, 1901.

27. Spady, A.A. and Kurbjun, M.C.: *Flight Management Research Utilizing an Oculometer,* SAE Internation Air Transportation Meeting, Boston, MA, 1978.

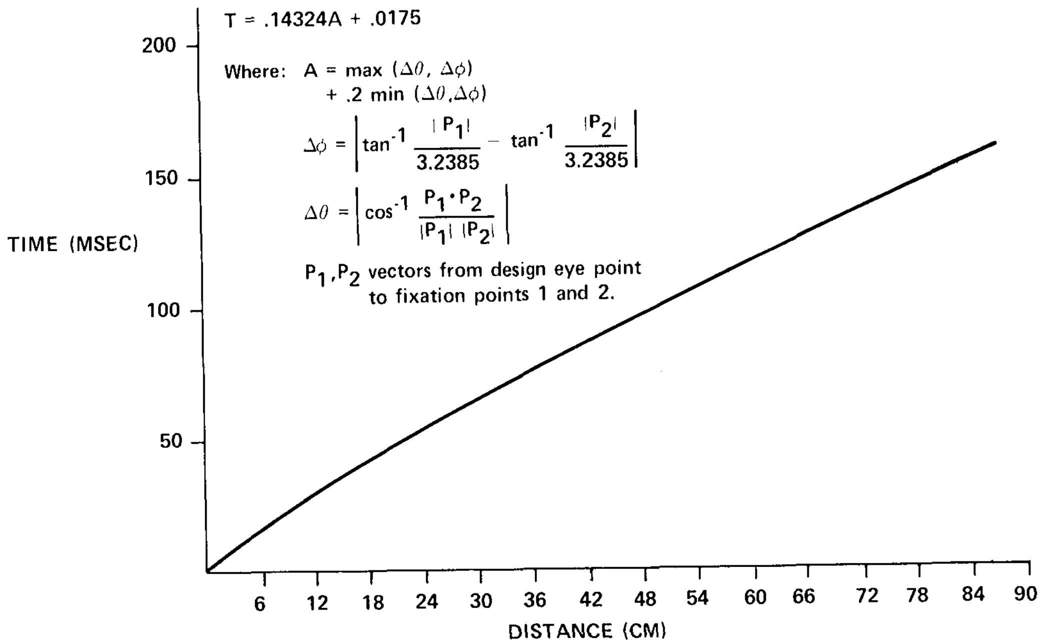


Figure 1.- Time for lateral eye movements as a function of distance between two targets.

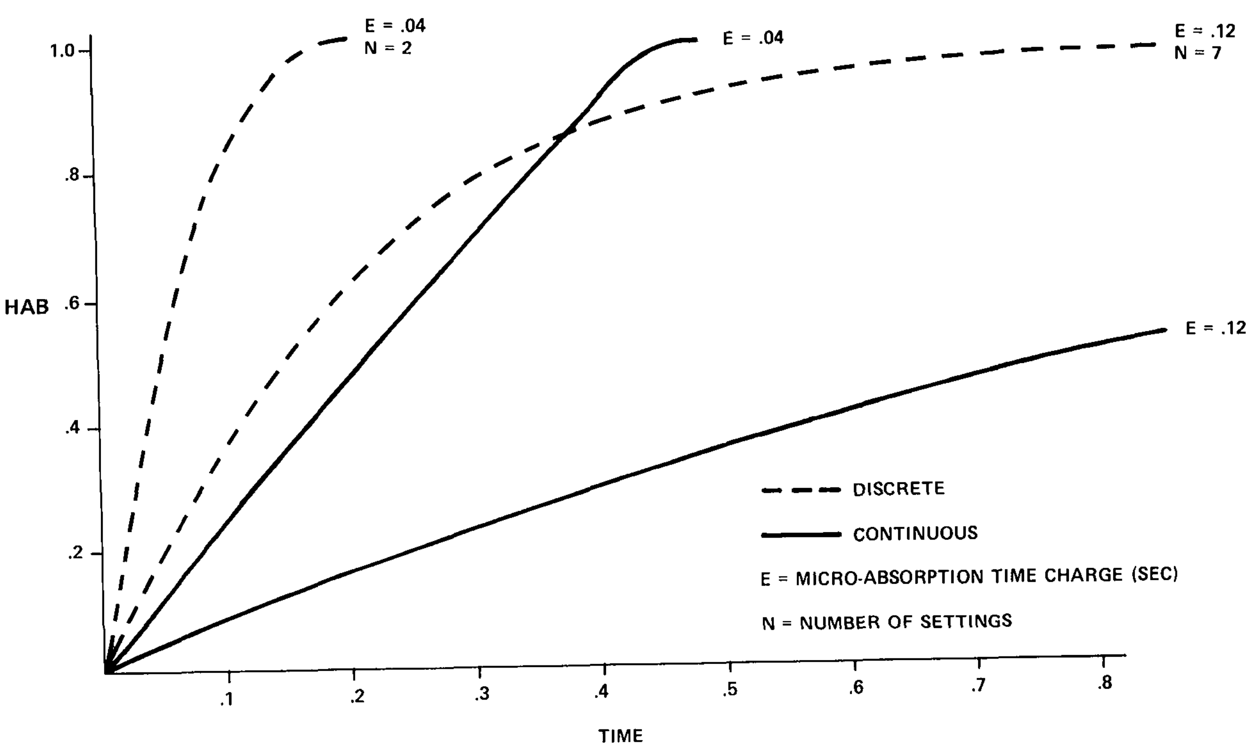


Figure 2.- Hab strength as a function of absorption time.

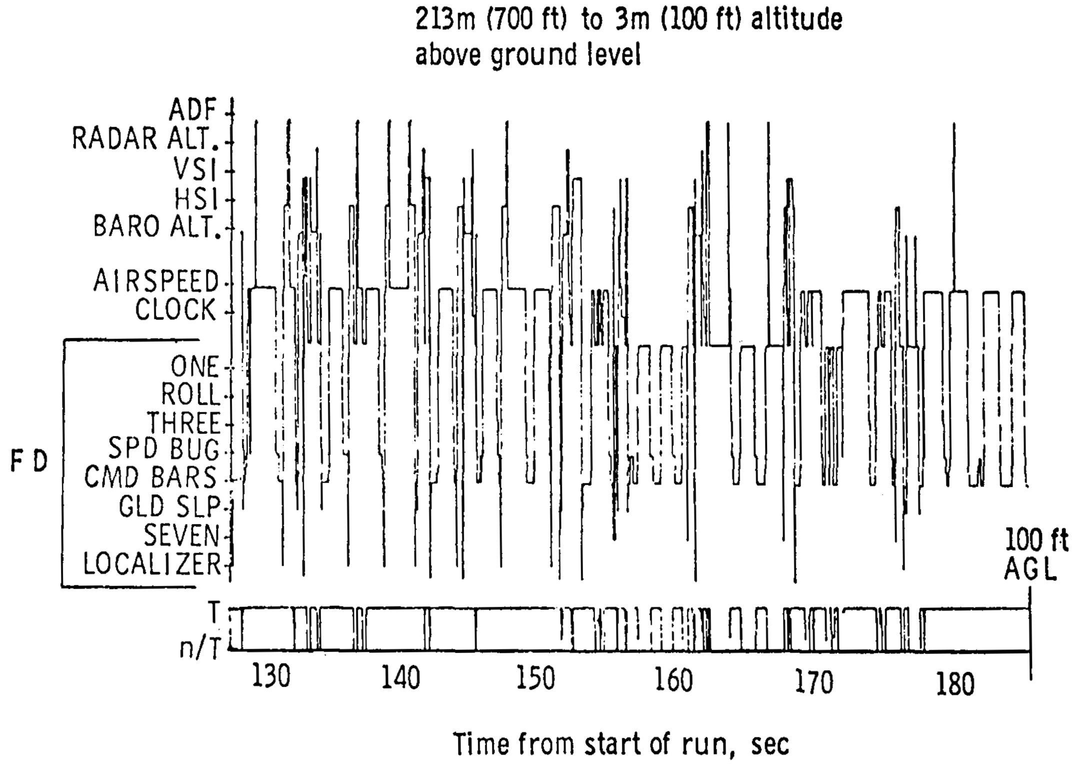


Figure 3.- Time history of one pilot's scan during coupled approach.

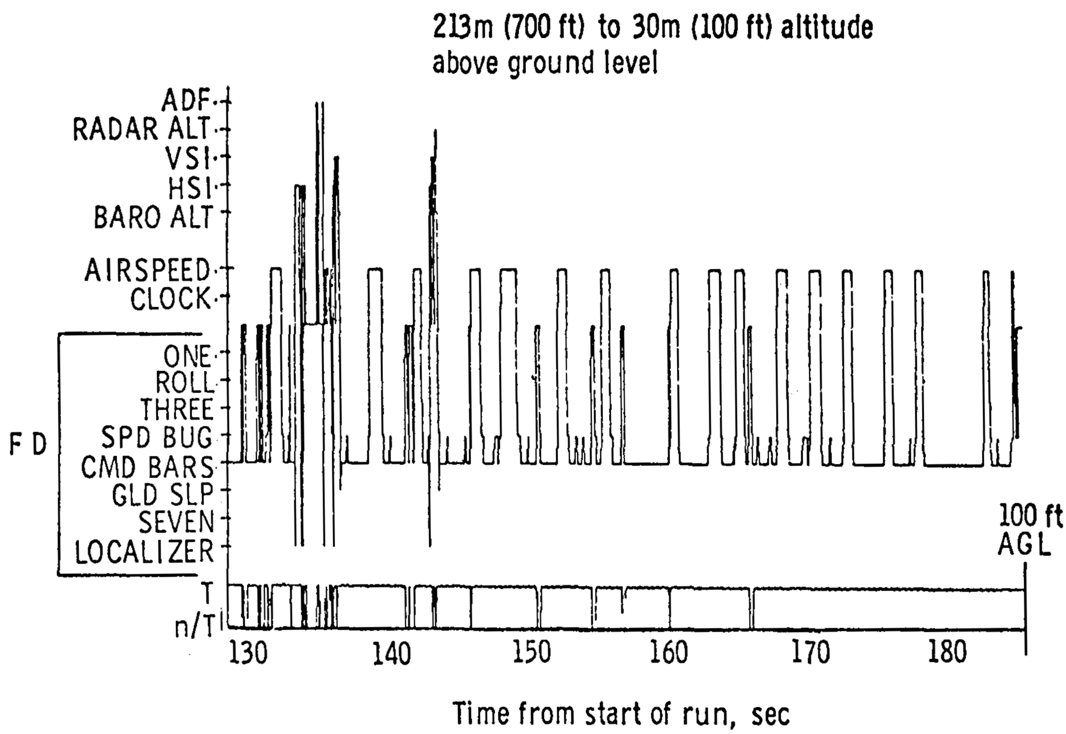


Figure 4.- Time histories of one pilot's scan during manual approach.

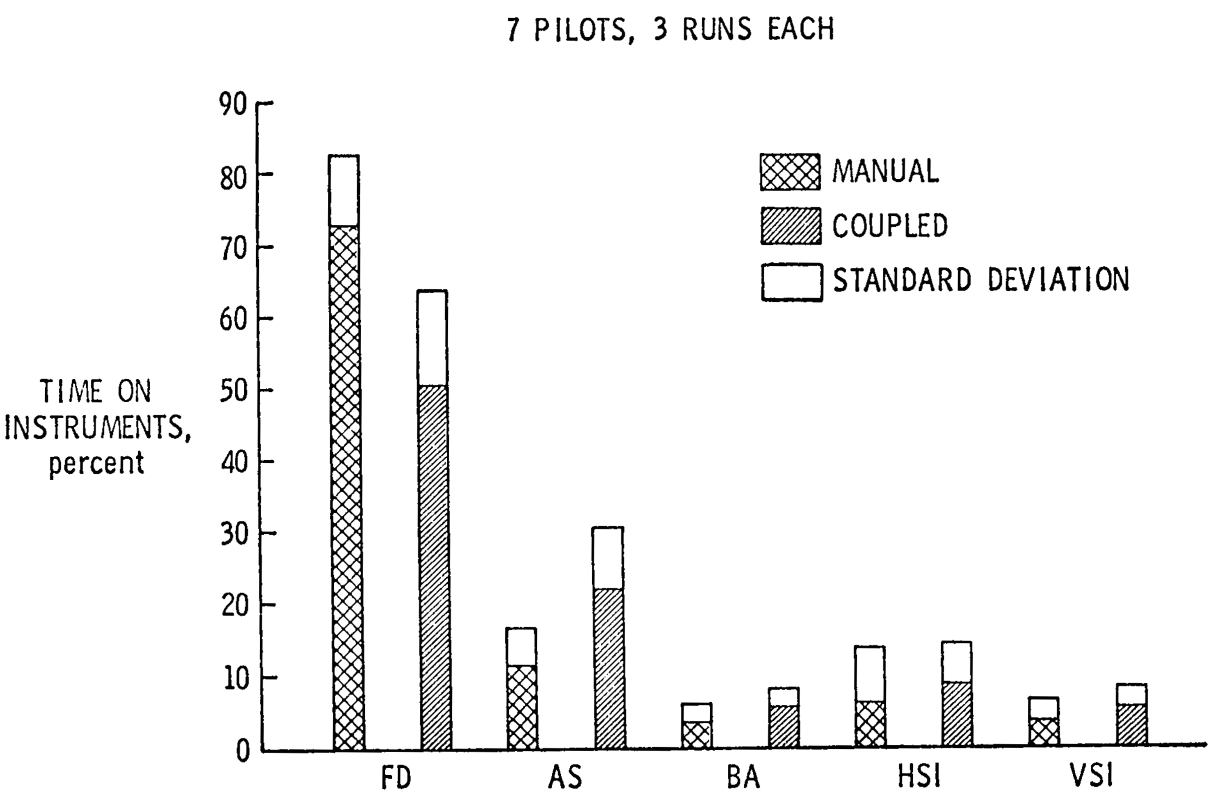


Figure 5.- Percent time on instruments for manual and coupled approaches.

A BRIEF OVERVIEW OF THE THEORY AND APPLICATION OF THE OPTIMAL CONTROL MODEL OF THE HUMAN OPERATOR

Sheldon Baron
Bolt Beranek and Newman Inc., Cambridge, Mass.

SUMMARY

This tutorial reviews the Optimal Control Model of the human operator. First, underlying motivation and concepts are presented, along with a review of the development and application of the model. Then, the structure of the model is described. Finally, results validating the model are presented.

INTRODUCTION

This paper reviews the Optimal Control Model (OCM) of the human operator developed principally by Kleinman, Levison, and the author (refs. 1 and 2, for example) at Bolt Beranek and Newman Inc. The OCM was originally developed for describing and predicting total system performance in continuous, manual control tasks. However, the model (or portions of it) has proven to be useful in a broader range of problems. Moreover, though not intended to be a structural analog of the human operator, many features of the model have interesting interpretations from an information processing view of human performance (ref. 3). The aim of this paper is to provide the reader with an overview of the OCM and a guide to the literature for more detailed information. Accordingly, it begins with a discussion of underlying motivation and a review of the development and application of the model. This is followed by a discussion of the important structural features of the model, some basic validation results and brief concluding remarks.

MOTIVATION AND REVIEW

The human controller is self-adaptive and, if motivated and given information about his performance, will attempt to change characteristics so as to perform better. On the other hand, human performance is limited by certain inherent constraints or limitations and by the extent to which the human understands the objectives of the task. These observations serve as the basis for the fundamental assumption underlying the OCM, namely, that the well-motivated, well-trained human operator will act in a near optimal manner subject to the operator's internal limitations and understanding of the task. This assumption is not new in manual control (e.g., (ref. 4)) or in traditional human engineering (e.g., Simon (ref. 5) calls it the Principle of Bounded Rationality). What is novel are the methods used to represent human limitations, the inclusion in the model of elements that compensate optimally for these limitations, and the extensive use of state-space concepts and the techniques of modern control theory.

Clearly, if the basic optimality assumption is to yield good results, it is necessary to have reliable, accurate, and meaningful models for human limitations. Insofar as possible, these models (or their parameters) should reflect intrinsic human limitations or should depend primarily on the interaction of the operator with the environment and not on the specifics of the control task. It is also desirable that the description of human limitations involves as few parameters as possible and that it be commensurate with the modern control system framework that is being employed. These principles have guided the development of the models for human limitations that will be described below.

There were several reasons for employing a modern control approach to analyzing manual control tasks, even though methods based on classical control techniques had been fairly successful. Initially, the principal motivation was provided by the basic logic of the optimality assumption and by the belief that state-space techniques provided a systematic approach to multi-input, multi-output systems that avoided some of the difficulties associated with the application of multi-loop analysis to man-in-the-loop problems. The powerful computational schemes associated with these techniques also were attractive in light of the complex monitoring and control problems that were becoming of interest. The basic approach to human limitations and the optimality assumption appeared to suggest a model that might adapt to task specifications and requirements "automatically" and not through a subsidiary set of adjustment rules. Finally, it was expected that the use of a normative model[1] and time-domain analysis would facilitate "modular" and "graceful" development of the model as new facets of human behavior were considered and understood.

A review of the progress and evolution of the OCM will provide some feel for the extent that the above-mentioned objectives and expectations have been fulfilled. Further insights will be provided by the discussions of the model and the validation results.

---

[1]The model is normative in that it predicts what the human should do, given his limitations and the task. Thus, for a new situation, one need only determine the operative limitations and what should be done. The fact that this assumption works well is testimony to the adaptability and capability of the trained human operator.

The first large-scale attempt at using the machinery of optimal control theory to model the human controller was initiated by Elkind et al. (ref. 6). Their study demonstrated the feasibility of predicting control characteristics and display requirements by systems analysis techniques based on optimal control theory. However, extremely simple versions of the human's limitations, information processing behavior, and compensation were used, leading to gaps and deficiencies in the results. What is essentially the current structure of the OCM was first proposed by Baron and Kleinman (ref. 1). They also proposed a visual scanning model that could be included in the optimization framework. Levison, Baron, and Kleinman (ref. 7) established the connection between observation noise and controller remnant, thus relating a measurable human limitation to parameters of the OCM and providing a mechanism for predicting remnant. Baron, Kleinman, et al. (ref. 8) used the remnant results and the structure developed previously to predict human performance in a complex, multi-loop VTOL hover task. These results demonstrated that one could proceed from relatively simple calibration experiments on single displays to prediction and explanation of human behavior in more realistic tasks involving two displays. This study also revealed the importance of including bandwidth limitations and randomness (motor-noise) at the controller's output as part of representation of human limitations.

Kleinman, Baron, and Levison (ref. 2) showed that the model could be used with a relatively invariant set of parameters quantifying human limitations to predict performance in three basic tracking tasks involving a range of control strategies. Excellent agreement between experimental data and model predictions of describing functions, remnant spectra, and state and control variances was obtained. This provided the most detailed validation of the model and demonstrated its capability for adapting to different control situations without resorting to auxiliary adjustment rules.

Baron and Kleinman (ref. 9) applied the model to study the human's precision control of a hovering VTOL-type vehicle. The effects of changes in aircraft stability derivatives on rms hovering performance were computed using the model. The results were compared with experimental simulator data and showed excellent correlation (within $\pm 1\,\sigma$ in the data) in most cases. In this study, parameters characterizing the pilot were essentially the same as for the basic tracking tasks mentioned above.

Kleinman and Baron (ref. 10) analyzed a piloted approach-to-landing task to evaluate pictorial display requirements. This problem involved a time-varying information base for the pilot. The effects of different display formats and display symbology were predicted in cases where the aircraft was subjected to turbulence and/or constant updrafts. The ability of the pilot to estimate these external disturbances and take the appropriate corrective action to minimize glide path errors was analyzed. Predictions of system performance were compared with data obtained in independent experimental investigations. The model-data agreements were excellent and demonstrated the model's ability to predict the time-varying adaptability of a pilot to updraft disturbances. In addition, the agreement between model results and data for cases in which there was no turbulence disturbing the aircraft provided further evidence of the validity of the model for human randomness (remnant).

Theoretical and empirical work proceeded to extend the model to more realistic situations and more complex systems. Levison et al. (ref. 11) developed and tested a mechanism for predicting task-interference in multi-task environments (not involving scanning). In addition, a method for estimating the relative attentional workload associated with a given task was devised. Levison (ref. 12) also investigated the relationship between observation noise and certain display characteristics. This provided direct empirical evidence for the scaling observation noise model and also showed how an equivalent observation noise could be used to account for perceptual thresholds. Levison and Kleinman (ref. 13) modeled a carrier-approach task that involved varying display gains, sudden changes in information base, and a more complex time-varying disturbance. Baron and Levison (ref. 14) used the model as a basis for a display analysis methodology and applied it to the analysis of vertical situation displays for STOL. The response to wind shears and the design of flight directors were also considered. These latter two studies were analytic in nature and did not involve any experimental verification.

Kleinman and Killingsworth (ref. 15) used the OCM to predict pilot performance during the flare and touchdown phase of STOL aircraft landing. This was an ambitious modelling effort since the vehicle dynamics were highly complex, ground effects and turbulence affected the motion of the aircraft, and the pilot was required to land within a short touchdown area. To analyze this situation, the model was extended to include the generation of open-loop commands by the human operator. In this study, model predictions were made first; subsequent comparison of these results with the test data showed very good agreement.

Kleinman and Perkins (ref. 16) used the OCM in an antiaircraft tracking task. The operator's task was to track an aircraft target in both azimuth and elevation using a visual gunsight. The dynamics of the sight and associated gun mount varied with time, making the tracking task very difficult. In addition, the target motion could be quite arbitrary (although not stochastic) and was unknown a priori by the gunner. Comparison of model vs. human ensemble statistics for the several typical aircraft trajectories showed good qualitative and quantitative agreement. Baron and Levison (ref. 17) also applied the OCM to data obtained from a simulated antiaircraft tracking task. This application demonstrated the model's utility in analysis and interpretation of experimental data. In particular, it showed that parameters of the perceptual portion of the OCM were affected in consistent ways by manipulation of experimental variables related to visual processing.

Harvey and Dillow (ref. 18) applied the OCM to predict pilot performance in air-to-air combat. They reported that "The major conclusion is that the model worked!" and that it was "reasonably simple to develop." Significantly, they used model parameters which, with the exception of motor noise, corresponded to those used in previous applications of the OCM.

The model was also being used to develop systematic design procedures for systems involving closed-loop control. As noted above, Baron and Levison (ref. 14) proposed a display design methodology based on the OCM. This methodology utilized performance/workload tradeoffs generated by the OCM to arrive at information requirements and certain display requirements to meet system specifications. Similar ideas were utilized to analyze both display and control characteristics for an aircraft with an advanced avionics configuration (ref. 19). Hess (ref. 20) proposed a more formal display design procedure using the OCM and included predictions of pilot rating as part of the process. Hoffman, Curry, et al. (ref. 21) developed a methodology aimed at display design for highly automated aircraft. They examined problems of simultaneous monitoring and control and explored different metrics for monitoring performance and workload with the aim of developing techniques for investigating tradeoffs between control and display sophistication.

Although display problems have received the most attention, other aspects of the system design problem have not been neglected completely. Levison (ref. 22) has explored the use of the model in analyzing control stick design problems in a vibration environment. Stengel and Broussard (ref. 23) have used the basic structure of the OCM, along with some assumptions concerning suboptimal adaptation, to determine stability boundaries in high-g maneuvering flight. And, recently, Schmidt (ref. 24) has proposed a design procedure for stability augmentation systems based on closed-loop analysis with the OCM.

The increased interest in flight simulators has spurred some additional extensions and applications of the model. Grunwald and Merhav (ref. 25) and Wewerinke (ref. 26) have incorporated mechanisms for describing the utilization of external visual cues in the OCM and have obtained preliminary experimental validation of their approaches. Although the subtleties and complexities associated with human perception of a complex scene are by no means resolved, these studies do suggest that the OCM could be useful for analyzing closed-loop control behavior based on external visual cues. The OCM has also been used to model continuous control performance in a multi-cue environment. Levison and Junker (ref. 27) studied roll-axis tracking in disturbance-regulation and target-following tasks and compared performance when only visual cues were available with performance when the visual cues were augmented with confirming motion cues. They found that the OCM could provide a task-independent framework for explaining performance under all possible experimental conditions. The availability of motion cues was modelled by augmenting the set of perceptual variables to include position, rate, acceleration, and acceleration rate of the motion simulator. This straightforward informational model allowed accurate model predictions of the effects of motion cues on a variety of response measures, for both the target-following and disturbance-regulation tasks.

In a somewhat different vein, Baron, Muralidharan, and Kleinman (ref. 28) used the OCM to develop a closed-loop model for analyzing engineering requirements for flight simulators. They predicted the effects on performance of certain simulation design parameters, such as an integration scheme and a sample rate. Model predictions were later verified in an empirical study by Ashworth et al. (ref. 29).

The above studies all focused on the operator in continuous control tasks. But the structure of the OCM, particularly the information processing submodel, also lends itself to modelling tasks in which monitoring and decision-making are the major concerns of the operator. The first attempt to exploit this aspect of the OCM was by Levison and Tanner (ref. 30) who studied the problem of how well subjects could determine whether a signal, embedded in added noise, was within specified tolerances. Their experiments were a visual analog of classical signal detection experiments except that "signal-present" corresponded to the situation of the signal being within tolerance. They retained the estimator/predictor and the equivalent perceptual models of the OCM and replaced the control law with an optimal (Bayesian) decision rule just as has been used in some popular behavioral decision-theory models. Model predictions compared favorably with experimental data for a variety of conditions involving different signal/noise ratios and different noise bandwidths.

Phatak and Kleinman (ref. 31) examined the application of the OCM information processing structure to failure detection and suggested several possible theoretical approaches to the problem. Gai and Curry (refs. 32 and 33) used the OCM information processing structure to analyze failure detection in a simple laboratory task and in an experiment simulating pilot monitoring of an automatic approach. They reported good agreement between predicted and observed detection times for both the simple and more realistic situations. In the latter case, the model was used in a multi-instrument monitoring task and accounted for attention sharing in the usual OCM fashion.

Finally, as indicative of future directions for OCM research, a recent study of Muralidharan and Baron (ref. 34) should be mentioned. In this work, the information processing structure of the OCM was used in conjunction with control and decision theoretic ideas to model ground-based operator control of a number of remotely piloted vehicles. Though the results have not been subjected to experimental validation, they demonstrate that these techniques are suited to the analysis of systems in which operators make decisions at discrete times and exercise direct control infrequently. In other words, the techniques appear suitable for supervisory control problems.

MODEL DESCRIPTION

In this section, the detailed structure of the OCM is reviewed. The discussion will be conceptual and verbal; the reader is referred to the previous references, particularly references 2 and 8, for mathematical details. Also, some relations to more traditional human performance theories will be mentioned.

In order to apply the OCM, the following features of the environment must be given: 1) a linearized state variable representation or model of the system being controlled; 2) a stochastic or deterministic representation of the driving function or environmental disturbances over which the operator must exert control; 3) a linearized "display vector" summarizing the sensory information utilized by the operator (including visual, vestibular, and other sources as appropriate); and 4) a quantitative statement of the criterion or performance index for assessing operator/machine performance. Criteria such as minimizing rms tracking error and control effort are typical. The specific assumptions concerning this description that are necessary to apply the theory are given in reference 2.

Given this environmental description, the model of the operator's behavior incorporates the elements shown in Figure 1. The figure illustrates only a single dimensional control task but the variables illustrated should be regarded as multi-dimensional vectors. First, the displayed variables are assumed to be corrupted by "observational noise" introduced by the human operator.[2] This noise is analogous to the internal noise level postulated in signal detection theory and provides one means by which the model can mimic human limitations in processing and attentional capacity. Different noise levels may be assumed for different displayed variables, and, if several visual displays are providing useful information, the noise level associated with each may be adjusted to account for the distribution of attention assigned by the operator. Alternatively, a model of attentional scanning (ref. 11) may be introduced to predict the noise level associated with each variable in order to produce optimal performance with respect to the criterion variable. This attention sharing model is crucial for predicting performance in complex, multivariable tasks. It can also serve as a basis for developing a variety of operator monitoring models (ref. 35).

At this point the model is dealing with a noisy representation of the displayed quantities. That representation is then delayed by an amount, τ, representing internal human processing delays. It is possible to assume differential delays for different sensory channels, but this additional complication has not been found necessary in past model applications to manual control data.

---

[2] If visual or indifference thresholds are important, such as with nonideal displays or external visual cues, these can be introduced in the model at this point (ref. 10). The method employed involves a statistical threshold that results in a rapid increase in observation noise when the signal is below the assumed threshold value. This is directly analogous to the threshold notions of signal detection theory.

The central elements of the model are represented in the blocks described as the Kalman estimator and predictor. Their purpose is to generate the best estimate of the current state of the displayed variables, based on the noisy, delayed perceptual information available. These blocks compute the estimate of this state so as to minimize the residual estimation uncertainty. What is being captured is a representation of the operator's ability to construct, from his understanding of the system and his incomplete knowledge of the moment-by-moment state of the system, a set of expectancies concerning the system behavior at the next moment in time. It is in these blocks that it is assumed that the operator has both an internal model of the dynamics of the system being controlled and a representation of the statistics of the disturbances driving the system. This representation is analogous to the schema of current human performance theories, and it is interesting to note that, in this formulation, the schema must incorporate knowledge of both the expected signals and the system dynamics being controlled.

Given the best estimate of the current system state, the next block assigns a set of control gains or weighting factors to the elements of the estimated state in order to produce control actions that will minimize the defined performance criterion. As might be expected, the particular choice of the performance criterion determines the weighting factors and thus the effective control law gains.

Just as an observation noise is postulated to account for input processing inadequacies, a motor noise is introduced to account for an inability to generate noise-free output control actions. In many applications this noise level is insignificant in comparison to the observation noise, but where very precise control is important to the conditions being analyzed, motor noise can assume greater significance in the model. Finally, the noisy output is assumed to be filtered or smoothed by a filter that accounts for an operator bandwidth constraint. In the model, this constraint arises directly as a result of a penalty on excessive control rates introduced into the performance criterion. The constraint may mimic actual physiological constraints of the neuromotor system or it may reflect subjective limitations imposed by the operator.

As the previous discussion shows, control strategy and motor response are separated from information processing in the OCM. This structure allows the OCM to be modified so as to treat decision-making problems. The estimator/predictor portion of the model generates all the statistical information necessary for optimal decision-making, given the assumptions that have been made concerning the system. Thus, by simply replacing the control law with an appropriate decision rule, one has a theoretical model for human decision making. For a normative model, the decision rule must be determined from optimization of an appropriate decision criterion (such as expected utility).

This, then, provides a conceptual description of the elements of the Optimal Control Model of the human operator. It should be emphasized that the parameter values that must be provided by the investigator correspond to the human limitations that constrain behavior. With these limitations as the constraints within which performance is produced, the model predicts the best that the operator can do. A large backlog of empirical research provides the data necessary to make realistic estimates of the appropriate parameter settings in the manual control context. This research has shown that these parameters are relatively invariant with respect to changes in task environment, thus enhancing the model's predictive capacity.

OCM VALIDATION STUDIES

The Optimal Control Model has been validated against experimental data for a variety of tasks, and detailed results may be found in the previously cited references. Here, a few of these results are presented in order to provide the reader with more of the background and with some feeling for the modelling accuracy attainable with the OCM.

Figures 2 and 3 (from ref. 2) illustrate the model's validity for two simple, but important systems: rate (K/s) and acceleration (K/s^2) command systems. In the figures, measured and theoretical human controller describing functions (h_e) and remnant spectra (Φ_{rr}) are compared. The describing function gain and phase may be thought of as measures of control strategy, whereas the remnant may be considered a measure of operator randomness. As can be seen, the model reproduces the characteristics of the subjects with remarkable fidelity. Moreover, the parameters of the model that quantify pilot limitations are virtually constant for the two situations. Table 1 compares measured and theoretical scores for the above cases. Results for a position command (K) system and for two tasks involving attitude regulation of a high performance aircraft are also shown. It is important to note that these results were obtained with a highly constant, though not identical, set of parameter values. (See ref. 36.)

These early single-input single-output studies served as the basic means of validating the model, but the OCM was principally directed at modelling human performance in more complicated situations. As we have discussed, an important part of this modelling is accounting for attention-sharing on the part of the operator. The basic empirical validation for the attention-sharing model was obtained in a four-axis tracking task (ref. 11). In this task, subjects had to control four independent rate-control systems with the errors in each system presented on separated displays. The subjects were required to fixate one display and use peripheral vision for tracking the other axes throughout the experiment (i.e., scanning was not allowed). The results for each axis performed alone and for all four together are presented in Table 2. Again, theoretical and measured results are in close agreement. Note that the effect of interference on total score is predicted better than its effect on individual scores. This appears to be true in other tests of the interference model, too. Analytic investigations of the tasks show that, for these experiments, tradeoffs in performance between subtasks do not effect overall performance substantially. When this is the case, the subjects are not motivated to seek the "absolute" optimal allocation (and they may not obtain the necessary feedback in training). Then, idiosynchratic behavior becomes more acceptable. The effects of attention sharing on the operator's describing function and remnant are given in reference 16. The result of adding a task is an increase in remnant, a decrease in operator gain, and an increase in high frequency phase lag. All these effects are predicted quite accurately by the OCM and the attention-sharing model.

CONCLUDING REMARKS

To summarize, the OCM has proven capable of predicting or matching human performance with considerable fidelity in a variety of tasks. Model parameters that account for basic human limitations have been isolated and shown to be essentially independent of system dynamics and forcing function characteristics; this enhances the model's predictive capability. Furthermore, submodels and parameters that reflect changes in display characteristics (such as thresholds, multiple displays, etc.) have been developed. An advantage of the OCM is that it contains an explicit model for information processing that also allows it to be used for analyzing monitoring and decision-making behavior.

There are, of course, limitations and problems associated with the model and its application. A major problem is the selection of an appropriate performance index in complex, realistic tasks. Though fairly systematic methods exist for making this selection, there is no guarantee that human operators will optimize the criterion selected by the theorist rather than some other, subjective one. Another limitation is the assumption of a perfect internal model. While this works quite well for trained operators, it can cause problems in modeling the performance of naive subjects (such as those in training) and can increase computational complexity beyond that which is necessary.

REFERENCES

1. Baron, S. & Kleinman, D. L.: The Human as an Optimal Controller and Information Processor. NASA CR-1151, September 1968. (Also IEEE Trans. Man-Machine Systems, Vol. MMS-10, No. 1, March 1969, pp. 9-16.)

2. Kleinman, D. L., Baron, S., & Levison, W. H.: An Optimal-Control Model of Human Response, Part 1: Theory and Validation. Automatica, No. 6, 1970, pp. 357-369.

3. Pew, R. W. and Baron, S.: The Components of an Information Processing Theory of Skilled Performance Based on an Optimal Control Perspective. Information Processing in Motor Control and Learning, Academic Press, Inc., N.Y., 1978, pp. 71-78.

4. Roig, R. W.: A Comparison Between Human Operator and Optimum Linear Controller RMS-Error Performance. IRE Trans. on Human Factors in Electronics, HFE-3, 1962, pp. 18-22.

5. Simon, H. A.: Models of Man. John Wiley & Sons, New York, 1957.

6. Elkind, J. I., Falb, P. L., Kleinman, D. L., & Levison, W. H.: An Optimal Control Method for Predicting Control Characteristics and Display Requirements for Manned-Vehicle Systems. Wright-Patterson Air Force Base, AFFDL-TR-67-187, June 1968.

7. Levison, W. H., Baron, S., & Kleinman, D. L.: A Model for Human Controller Remnant. IEEE Trans. on Man-Machine Systems, MMS-10, No. 4, December 1969, pp. 101-107.

8. Baron, S., Kleinman, D. L., Miller, D. C., Levison, W. H., & Elkind, J. I.: Application of Optimal Control Theory to the Prediction of Human Performance in a Complex Task. Wright-Patterson Air Force Base, AFFDL-TR-69-81, March 1970.

9. Baron, S. & Kleinman, D. L.: Prediction and Analysis of Human Performance in a VTOL Hover Task. Proceedings of Seventh Annual NASA-University Conference on Manual Control, USC, Los Angeles, California, June 1971.

10. Kleinman, D. L. & Baron, S.: Analytic Evaluation of Display Requirements for Approach to Landing. NASA CR-1952, November 1971.

11. Levison, W. H., Elkind, J. I., & Ward, J. L.: Studies of Multi-Variable Manual Control Systems: A Model for Task Interference. NASA-Ames Research Center, NASA CR-1746, May 1971.

12. Levison, W. H.: The Effects of Display Gain and Signal Bandwidth on Human Controller Remnant. Wright-Patterson Air Force Base, AMRL-TR-70-93, March 1971.

25. Grunwald, A. J. and Merhav, S. J.: Vehicular Control by Visual Field Cues--Analytical Model and Experimental Validation. IEEE Trans. Systems, Man and Cybernetics, Vol. SMC-6, No. 12, March 1976, pp. 835-845.

26. Wewerinke, P. H.: A Theoretical and Experimental Analysis of the Outside World Perception Process. Proc. of Fourteenth Annual Conference on Manual Control, NASA Conference Publication 2060, November 1978, pp. 535-556.

27. Levison, W. H. and Junker, A.: A Model for the Pilot's Use of Motion Cues in Roll-Axis Tracking Tasks. Proc. of Thirteenth Annual Conference on Manual Control, MIT, June 1977, pp. 377-388.

28. Baron, S., Muralidharan, R. and Kleinman, D. L.: Closed-Loop Models for Analyzing Engineering Requirements for Simulators. BBN Report No. 3718, Bolt Beranek and Newman Inc., Cambridge, Mass., May 1978.

29. Ashworth, B. G., McKissick, B. T., and Parrish, R. V.: The Effects of Simulation Fidelity on Air-to-Air Tracking, Proc. of the Fifteenth Annual Conference on Manual Control, WPAFB, Ohio, March 1979.

30. Levison, W. H. and Tanner, R. B.: A Control-Theory Model for Human Decision Making. NASA CR-1953, December 1971.

31. Phatak, A. & Kleinman, D. L.: Current Status of Models for the Human Operator as a Controller and Decision Maker in Manned Aerospace Systems. Automation in Manned Aerospace Systems. AGARD Proc., Dayton, OH, No. 114, October 1972.

32. Gai, E. G. and Curry, R. E.: Failure Detection by Pilots During Automatic Landing; Models and Experiments. Proc. of Eleventh Annual Conference on Manual Control, NASA TMX-62, 464, May 1975, pp. 78-93.

33. Gai, E. G. and Curry, R. E.: A Model of the Human Observer in Failure Detection Tasks. IEEE Trans. on Systems, Man and Cybernetics, Vol. SMC-6, No. 2, February 1976, pp. 85-93.

34. Muralidharan, R. and Baron, S.: Combined Monitoring, Decision and Control Model for the Human Operator in a Command and Control Task. Proc. Fourteenth Annual Conference on Manual Control, NASA Publication No. 2060, November 1978, pp. 647-662.

35. Kleinman, D. L. and Curry, R. E.: Some New Control Theoretic Models for Human Operator Display Monitoring. IEEE Trans. on Systems, Man and Cybernetics, Vol. SMC-7, No. 11, November 1977, pp. 778-784.

36. Baron, S.: A Model for Human Control and Monitoring Based on Modern Control Theory. Journal of Cybernetics and Information Sciences, Vol. 4, No. 1, Spring 1976, pp. 3-18.

13. Levison, W. H. & Kleinman, D. L.: Analysis of Pilot/System Performance in Carrier Approach. BBN Report No. 2169, Bolt Beranek and Newman, Cambridge, Mass., September 1971.

14. Baron, S. & Levison, W. H.: A Display Evaluation Methodology with Application to Vertical Situation Displays for STOL Aircraft. Proceedings of Ninth Annual NASA-University Conference on Manual Control, M.I.T., May 1973, pp. 121-132.

15. Kleinman, D. L. & Killingsworth, W. R.: A Predictive Pilot Model for STOL Aircraft Landing. NASA CR-2374, March 1974.

16. Kleinman, D. L. & Perkins, T.: Modelling the Human in a Time-Varying Anti-Aircraft Tracking Loop. IEEE Trans. on Auto. Control, AC-19, 1974, pp. 297-306.

17. Baron, S. & Levison, W. H.: Analysis and Modelling Human Performance in AAA Tracking. BBN Report No. 2557, Bolt Beranek and Newman Inc., Cambridge, Mass., March 1974.

18. Harvey, T. R. & Dillow, J. D.: Fly and Fight: Predicting Piloted Performance in Air-to-Air Combat. Paper presented at the Tenth Annual Conference on Manual Control, Wright-Patterson Air Force Base, Ohio, 1974, pp. 625-640.

19. Levison, W. H. and Baron, S.: Analytic and Experimental Evaluation of Display and Control Concepts for a Terminal Configured Vehicle. BBN Report No. 3270, Cambridge, Mass., July 1976.

20. Hess, R. A.: Analytical Display Design for Flight Tasks Conducted Under Instrument Meteorological Conditions. IEEE Trans. on Systems Man and Cybernetics, SMC-7, No. 6, June 1977, pp. 453-461.

21. Hoffman, W. C., Curry, R. E., et.al.: Display/Control Requirements for VTOL Aircraft. NASA CR-145026, August 1975.

22. Levison, W. H.: Guide for the Design of Control Sticks in Vibration Environments. BBN Report No. 2863, Bolt Beranek and Newman Inc., Cambridge, Mass. August 1974.

23. Stengel, R. F., and Broussard, J. R.: Prediction of Pilot-Aircraft Stability Boundaries and Performance Contours. IEEE Trans. on Systems, Man and Cybernetics, Vol. SMC-8, No. 5, May 1978, pp. 349-356.

24. Schmidt, D. K.: Optimal Flight Control Synthesis Via Pilot Modelling. Journal of Guidance and Control, Vol. 2, No. 4, July-August 1979, pp. 308-312.

TABLE 1.- MEASURED AND THEORETICAL HUMAN PERFORMANCE

System	MS Error Meas.	Theor.	MS Control Meas.	Theor.
Position Control	.13	.14	.53	.54
Rate Control	.13	.12	4.2	3.83
Acceleration Control	.014	.014	1.43	1.28
High Performance Aircraft (Pitch)	.026	.026	.0032	.0034
High Performance Aircraft (Roll)	.03	.026	.080	.086

TABLE 2.- COMPARISON OF MEASURED AND PREDICTED ERROR VARIANCE

SCORES FOR 4-AXIS EXPERIMENT

	Viewing Condition					
	Measurement	Foveal	16° Periph Ref Ext.	16° Periph No Ref Ext	22° Periph No Ref Ext	Total Score
(a) Measured	1-axis	.11	.25	.42	.96	1.7
	4-axis	.27	.94	1.3	1.6	4.1
(b) Predicted:	1-axis	.11	.25	.39	.98	1.7
Optimal Behavior	4-axis	.49	.82	1.1	1.8	4.2

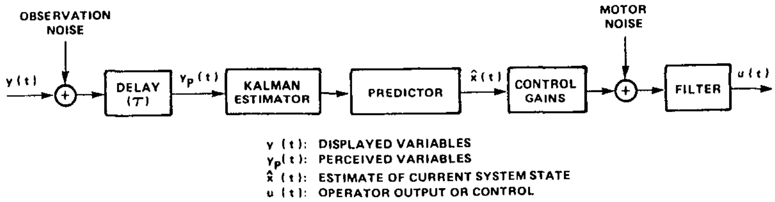


Figure 1.- Structure of OCM operator model.

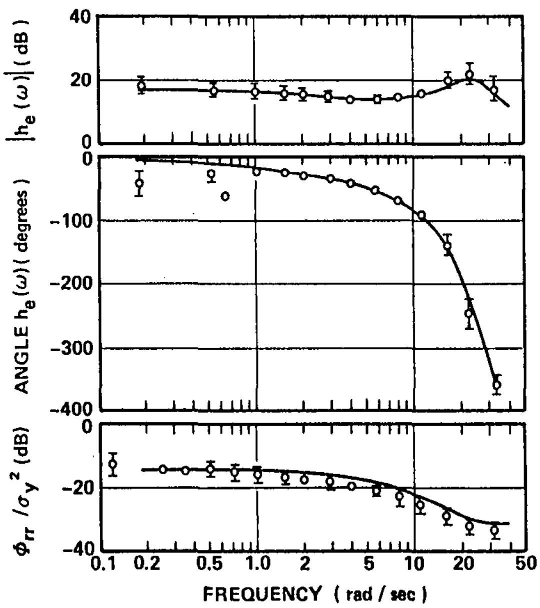


Figure 2.- Operator response - K/s dynamics.

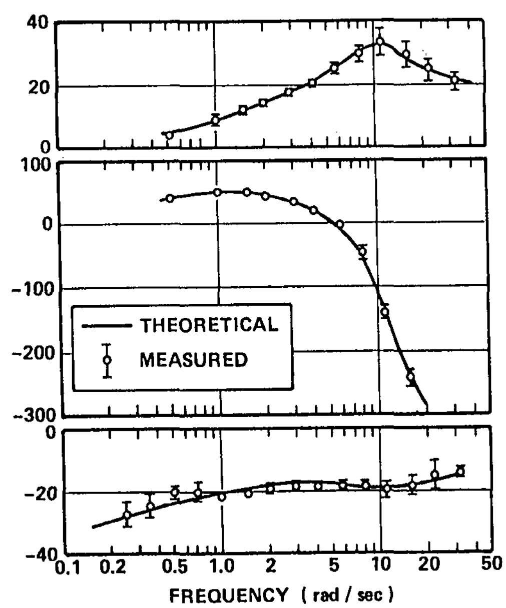


Figure 3.- Operator response - K/s^2 dynamics.

SAINT: A COMBINED SIMULATION LANGUAGE FOR MODELING MAN-MACHINE SYSTEMS

Deborah J. Seifert
Department of Operational Planning
AiResearch Manufacturing Company of Arizona

SUMMARY

SAINT is an acronym for: Systems Analysis of Integrated Networks of Tasks. SAINT is a network modeling and simulation technique for design and analysis of complex man-machine systems. SAINT provides the conceptual framework for representing systems that consist of discrete task elements, continuous state variables, and interactions between them. It provides a mechanism for combining human performance models and dynamic system behaviors in a single modeling structure. SAINT facilitates an assessment of the contribution that system components make to overall system performance.

INTRODUCTION

SAINT is a computer simulation tool for modeling and analyzing man-machine systems. While SAINT was designed for modeling manned systems in which human performance is a major concern, it is potentially applicable to to a broad class of problems--those in which discrete and continuous elements are to be portrayed and the behavior exhibits time varying properties. SAINT provides a mechanism for describing these dynamics so analyses can be performed.

SAINT evolved from two separate technologies. Task analysis and the Monte Carlo simulation of operator performance under workload stress as represented by Siegel and Wolf (ref. 1) were the origin for the human factors development. Many of the features eventually incorporated into SAINT were identified as requirements based upon experience in applying this technology. The second origin of SAINT was in the GASP family of simulation techniques (ref. 2). The earliest version of SAINT was an incorporation of the Siegel-Wolf model in a modified P-GERT package (ref. 3). The subsequent evolution of SAINT adapted features of GASP IV and allowed SAINT to become a flexible, sophisticated, combined modeling technique where networks of discrete events could be modeled along with the dynamics of continuous processes.

It is this ability to combine models of dynamics (e.g., aircraft equations of motion) with models of discrete activity sequences (e.g., operator actions) that permits the systems analyst to describe both hardware and human performance in the context of a single model. This affords the system engineer the opportunity to analyze system effectiveness and quantify the relative contributions of man and machine.

SAINT CONCEPTS

For the discrete event simulation, a graphical-network approach to modeling is taken, whereby a user of SAINT describes the system to be analyzed via a network model and auxiliary descriptions (e.g., equipment and operator performance parameters). A symbol set has been devised for diagraming the discrete task network. The SAINT computer simulation program accepts a description of the network to be simulated and automatically performs an analysis to obtain statistical estimates of system performance. For the continuous process representation, the user is expected to provide FORTRAN statements of the relevant state equations to be solved. Mechanisms are provided for creating an interaction between the discrete and continuous components of the model.

Discrete Task-Oriented Model Component

The discrete task-oriented component of the SAINT model consists of nodes and branches, each node representing a task. Tasks are described by a set of characteristics (e.g., performance time duration, priority, resource requirements). Branches connecting the nodes indicate precedence relations and are used to model the sequencing and looping requirements among the tasks. Complex precedence relations have been designed into SAINT to allow predecessor-successor relationships which are deterministic, probabilistic, and conditional. Resources, either human operators or hardware equipment, perform the tasks in accordance with the network prescribed precedence relations, subject to resource availability. The precedence relations also indicate the flow of information through the network. Information is organized into packets, with each packet containing attributes that characterize the information being processed. The information packet can characterize items flowing through the network, or any other concept related to network flow. When a task is completed, the information packet residing at the task is transmitted along each precedence branch selected. Information attribute values can be assigned or modified at any task in the network and can influence both task performance times and task branching relations.

Resources perform tasks either individually or in groups. Each resource included in a SAINT model is described by a set of attributes. These attributes are also organized into packets, with each packet characterizing a particular resource. Examples of operator attributes include such parameters as level of training, age, height, etc. Machine reliability is an example of an equipment attribute. Resource attributes are used in conjunction with the task descriptions in order to make a general network model resource-specific. The initial values of these user-defined resource attributes are assigned prior to the start of the simulation. The values may be dynamically changed at any task in the network and can be used as parameters in determining both task performance times and precedence relations.

In many instances it may be desirable to specify attributes which are not directly applicable to an information-oriented or resouce-oriented characterization. These attributes are global in nature and do not flow through or move about the network as information and resource packets do. Temperature and time remaining in a mission are examples of model parameters which may be characterized as system attributes. Just as with information and resource attributes, system attributes may influence the task network performance and flow.

Each task in a SAINT network has two requirements which must be satisfied before the task can be performed. First, a specified number of predecessor tasks must be completed before the task is released. Second, once the task has been released, the resources required to perform the task must be available (that is, not be busy performing other tasks). All tasks which have been released (all predecessor requirements have been satisfied) but whose required resources are not available are ranked in a queue according to their priority. Task priority may be assigned at the start of the simulation and may change dynamically as a function of system parameters and contingencies. When the required resources become available with task completions, the tasks in the waiting queue are started. The time to perform a task may be specified as a random variable defined by a probability distribution. SAINT supplies the user with 11 different distributions (Normal, Gamma, Beta, Weibull, etc.)

Frequently the task performance time is also a function of the type of task, the resource or resources performing the task, the status of the system, or the condition of the environment at the time the task is executed. SAINT provides for the specification of factors which influence task performance via user-written moderator functions. It is presumed the modeler can describe (e.g., by least squares techniques) the functional relationshps between a set of conditions and a performance parameter or attribute of interest. For example, one might hypothesize that fatigue affects operator performance such that the average task time increases as a function of mission duration. Research data must be obtained to postulate the functional form of this relationship and fit a curve to these results. This empirically derived relationship can then be implemented in SAINT as a moderator function to determine the possible impact fatigue could have on operator performance. In addition to moderator functions, user-written functions can be developed for specifying attribute assignments. Both types of functions are written in FORTRAN or a FORTRAN-compatible language.

Contingencies, decision making, and emergency conditions can be represented via SAINT's flexible attribute assignment and branching logic. SAINT provides two additional mechanisms for modeling system performance.

The first of these is termed task modification. This feature enables the user to modify task parameters as a function of ongoing system events. For example, consider a task which may require repetition due to a possibility of failure on the first attempt. The second time the task is performed the performance time may be significantly smaller than the initial execution. SAINT provides for the modification of the task time distribution after the initial attempt. Other task parameters can be modified in a similar fashion.

The second SAINT modeling construct of interest is "clearing". Both tasks and resources can be cleared. "Task" clearing halts a specified task in progress, contingent on the completion of another task. "Resource" clearing halts whatever task the specified resource is performing. Both types of clearing may specify an additional task to be signaled. As an example, consider the simulation of an emergency condition in which all operators must stop their ongoing activites to assist in the emergency operations. This situation is best modeled in SAINT with resource clearing. The onset of the unexpected event would "free-up" (clear) the operators. Concurrently, emergency handling tasks would be signaled for initiation (and release if all other precedents were satisfied). Task and resource clearing provide dynamic realism in man-machine simulation modeling. The network symbol used to diagram a task in a SAINT model is illustrated in Figure 1. The input side of the node reflects the precedence requirements for releasing a task. The number of requirements for releasing a task the first time is on the top (PR1) and the number of requirements for releasing a task on subsequent times is on the bottom (PR2).

The center portion of the task symbol contains all task description information, such as performance time characteristics, statistics to be collected, and attributes to be assigned. It is subdivided into rows, with each row containing a specific type of descriptive information about the task. Further, each row is divided into two parts. The left-hand part contains the task description code. It is used to identify the type of information that appears in the right-hand part of the row, and can be any of the 17 available codes shown in Table I.

The LABL permits an eight character identifier to be associated with this node to depict the nature of the task/activity represented. The TIME parameters indicate the distribution type and parameter values for the characterization of task duration. If activity times are known to be a function of specifiable factors (e.g., task, system, or information attributes), a moderator function (MODF) may be employed (as a FORTRAN subroutine) to generate the activity duration instead of generating a time value by Monte Carlo methods. If Monte Carlo methods are employed (via TIME specification), a modification can be effected during model execution by using the DMOD feature to identify an alternate distribution and/or parameter set when specified event conditions prevail. RESR may be used to specify the type and quantity of resources and whether multiple resources imply substitution ("or") rather than conjoint ("and") requirements. If priority (PRTY) is a concern, it can be specified a priori and subsequently manipulated dynamically during model execution. Since information packets can arrive at a task from several sources, but only one will exit, it is necessary to specify which incoming packet will be passed along, INCM. The default condition for processing information packets is to simply pass the last one arriving at the node. If different predecessor completions are required in order for the task to be released, the DIFF option must be specified. Otherwise the multiple occurrence of any predecessor may cause the task to be prematurely released. When two or more tasks have identical completion times, it is necessary to specify which will take precedence (PREC) over the others. User-defined task characteristics (UTCH) permit the user to specify additional attributes of a task (e.g., difficulty, complexity, etc.), and these attributes can be modified upon task execution. Information, resource, and system attributes can be assigned or updated (ATAS) upon task release, start, or completion as required. The statistics to be collected (STAT) are described in subsequent discussion. A particular task can be used to mark the start point (MARK) for timing how long it takes to traverse a path to some other task of interest. The MARK feature allows elapsed time computations within the network (e.g., time between events). Task and resource clearing operations are established by specifying the appropriate parameters associated with the TCLR and RCLR mnemonics. Upon completion of a task, SWIT allows a switch or flag to be set for subsequent examination in the continuous state variable component of the model. The REGL mnemonic is used as a device for regulating values employed in the continuous process model, where a task is permitted to alter a state variable, for example.

By selectively using these description codes, only the information necessary to describe a task need be shown on the task symbol. In this manner, any or all of the task description codes can be specified for a particular task. If more than the four rows provided are required for a complete description, the user simply adds the necessary number of additional rows to the bottom of the task description portion of the task symbol.

The output side of the node contains the task number (TSK). In addition, the shape of the output side indicates the branching operation to be performed upon task completion. It specifies the process to be employed in selecting the successor tasks whose precedence requirements should be reduced by one. The four branching types included in SAINT are deterministic, probabilistic, conditional take-first, and conditional take-all. Their shapes are depicted in Figure 2.

When deterministic branching is specified, the number of requirements for all successor tasks is reduced by one. For probabilistic branching, each branch emanating from the task has an associated probability of selection. These probabilities may be specified directly or obtained from information, operator or system attributes. Only a single successor task is selected. For conditional take-first branching, each branch has an associated condition, and the branches are ordered. Each condition is tested in the prescribed order, and the first branch whose condition is satisfied is selected. Conditional take-all branching operates in the same manner, but selects all branches whose conditions are satisfied. Conditions may be based on task completions, simulated time, or attribute values.

The above discussion only included the basic task node symbology. Additional symbolism is available for task modification, task signaling as a result of task or resource clearing, task signaling resulting from a threshold crossing, and state variable monitors (refs. 4 and 5).

Continuous State Variable Model Component

The second component of a SAINT model is the state variable description. The SAINT user defines these state variables by writing the algebraic, difference, or differential equations that govern their time-dependent behavior. The use of state variables in SAINT is optional.

The SAINT user writes the state variable equations in a FORTRAN subroutine (subroutine STATE). State variables represented by algebraic or difference equations are defined in subroutine STATE as SS(•) variables. Those represented by differential equations are written in terms of DD(•) variables. SAINT employs a Runge-Kutta-England (RKE) numerical algorithm to integrate the equations of subroutine STATE written in terms of the DD(•) variables. The RKE algorithm obtains a solution to a set of simultaneous first order ordinary differential equations. Higher order differential equations can be modeled by placing the equations in canonical form. Subroutine STATE can be used to model state variables using a combination of DD(•) and SS(•) variables.

In SAINT, simulated time is advanced in accordance with the type of system being modeled. If no state variables are included, simulated time is advanced from one task completion to the next. When state variables are included in the model, time is also incremented in steps between scheduled task completions for the purpose of updating the values of the state variables. The step size is a function of user-specified accuracy requirements.

Discrete and Continuous Component Interactions

The interactions between tasks and state variables are initiated either by tasks being completed or by state variables crossing specified threshold values. Upon the completion of a task, state variables may be discretely regulated by increasing or decreasing their values. In addition, task completions can change the values of logical variables which can be used to alter state variable equation forms or the network structure. In this manner the discrete task-oriented component of the model affects the continuous state variable component.

Threshold crossings by state variables can signal or initiate tasks. Thus the values of state variables can influence task performance characteristics and precedence relations. Threshold crossings can also change the values of logical variables which, in turn, can be used to alter equation forms or change task precedence.

As an example of discrete and continuous component interactions, consider a system in which a pilot must keep the aircraft altitude within specified constraints. The pilot's inputs might be modeled as discrete tasks and the aircraft dynamics as continuous state variables. When the altitude state variable crosses the allowable threshold value, the corresponding discrete pilot makes the appropriate input and regulates the state variable(s) which determine altitude. Thus, through this component interaction, the aircraft altitude is brought back within acceptable limits.

STATISTICAL OUTPUT

Once the model has been built, the modeler can impose a data collection structure to obtain information about his description of the system as it is exercised. A variety of data can be obtained; these fall into four major categories. The first type of output is a statistical description of the execution of specific nodes or collections of nodes. There are sixteen possible combinations of interval and task completion statistics that one can collect using the built-in features of SAINT. Since users can create their own functions for updating attributes and for moderating network parameters, it has been necessary to allow the user to collect his own statistics on those parts of the model which cannot be predefined because the user creates them himself. SAINT supplies statistical subroutines for collecting data on user-supplied parts of the model. Tabular summaries of the computed descriptive statistics can then be generated to portray the results of a single iteration, a set of iterations, or a series of iterated runs showing the trends induced by some systematic variation of run conditions.

A second type of output which SAINT provides is resource utilization statistics. Information on the busy/idle status for both the human resources as well as the equipment resources is automatically presented at the completion of each simulation run. These statistics can be employed in evaluating workload and system capacity issues.

The third type of output is a graphic portrayal of the probability and cumulative density functions for a distributed variable. These histograms provide a quick look at the shape of the data. An experienced user can store the actual values on an external device; later, the data can be fed to a plotting package for reproducible drawings.

Time traces of the state variables are a fourth type of output. Up to 10 variables can be plotted on the same graph with user specified scale factors and plotting symbols. Multiple graphs can be generated. Tabled values of the variables can also be obtained. The tabulated plot provided by SAINT equips the user to quickly examine the results of his simulation run.

THE SAINT PROGRAM

Development of the SAINT simulation package has been completed and is fully documented (refs. 4, 5, 6, and 7). SAINT was developed in ANSI standard FORTRAN and, consequently, is machine-independent. The user, however, must supply his own system-specific random number generator. The task network data is punched on cards in free-form. SAINT includes an extensive input error-checking feature to assist users in debugging their models. For production runs, users can select a more efficient non-error-checking version of SAINT. A separate FORTRAN program has been devised to create a source module with the COMMON blocks sized to the problem being run. SAINT also includes provisions that allow formatting model outputs so they can be processed by available statistical analysis packages (ref. 7).

APPLICATIONS OF SAINT

SAINT has been used to analyze a wide variety of man-machine systems. It is gaining a wide and enthusiastic acceptance by systems modelers and analysts of many disciplines. The following is a list of completed or ongoing modeling and simulation efforts involving the use of SAINT: SAINT has been used by the Aerospace Medical Research Laboratory (AMRL) to evaluate alternatives for a Remotely Piloted Vehicle/Drone Control Facility (RPV/DCF) in which operators monitor and control the flight of RPV's through the use of visual (CRT) displays (ref. 8). SAINT was used, also, to provide flight control performance predictions for the Digital Avionics Information System (DAIS) cockpit configuration in which dedicated instruments, displays, and subsystem status displays have been replaced with interactive multipurpose displays and multifunction keyboard switching. A first model of this system employed discrete task networks to represent the pilot's activities and continuous state equations to represent the vehicle dynamics (ref. 9). More recently, a model of DAIS has been developed in which the pilot's discrete information storage and retrieval activities were modeled by tasks; however, the pilot's flight control was represented by a variation of the Optimal Control Model developed by Bolt, Beranek and Newman. In this combined discrete/continuous model of the human operator the pilot operates in a so-called "open loop" preprogrammed fashion between flight control variable sampling (ref. 10). SAINT is currently being used by AMRL to provide cost trade-off design analyses of proposed alternative configurations for the UPD-X All Weather Wide Area Surveillance ground exploitation station. AMRL plans to utilize SAINT to analyze design proposals in a B-52 strategic navigation system involving complex crew activities and task management (ref. 11). SAINT has been used by the Air Force Human Resources Laboratory to explore the feasibility of employing computer simulation for evaluating human effects on nuclear systems safety in a missile loading operation (ref. 12). SAINT was employed by Air Force Weapons Laboratory to examine workload sharing and nuclear radiation effects on pilot performance in an air-to-air refueling mission (ref. 13). Purdue University researchers utilized SAINT to investigate the effect of higher degrees of automation, different capacities of process limiting operations, and alternative task allocations on the operator's idle times in a hot strip mill (ref. 14). SAINT has been used by the U. S. Department of Commerce Office of Telecommunications to analyze communication frequency utilization in a railroad switching yard. SAINT has been used by New Mexico State to compare theoretical human performance predictions with empirically derived performance data (ref. 15). SAINT is being employed by Pritsker and Associates in support of the Army Research Institute to analyze human system performance in an AN/TSQ-73 guided missile air defense system operation (ref. 16). SAINT is also being utilized by several universities both in the classroom and for research activities. Among these are Purdue, Iowa State, North Carolina State, Ohio State, and Arizona State.

REFERENCES

1. Siegel, Arthur I.; and Wolf, J. Jay: Man-Machine Simulation Models: Performance and Psychosocial Interaction. New York, John Wiley & Sons, Inc., 1967.

2. Pritsker, A. A. B.: The GASP IV Simulation Language. New York, John Wiley & Sons, 1974.

3. Pritsker, A. A. B.; Wortman, D. B.; Seum, C. S.; Chubb, G. P.; and Seifert, D. J.: SAINT: Volume I, Systems Analysis of Integrated Networks of Tasks. Aerospace Medical Research Laboratory, Wright-Patterson Air Force Base, Ohio, AMRL-TR-78-126, 1974.

4. Wortman, D. B.; Duket, S. D.; Seifert, D. J.; Hann, R. L.; and Chubb, G. P.: Simulation Using SAINT: A User-Oriented Instruction Manual. Aerospace Medical Research Laboratory, Wright-Patterson Air Force Base, Ohio, AMRL-TR-77-61, 1978.

5. Wortman, D. B.; Duket, S. D.; Seifert, D. J.; Hann, R. L.; and Chubb, G. P.: The SAINT User's Manual. Aerospace Medical Research Laboratory, Wright-Patterson Air Force Base, Ohio, AMRL-TR-77-62, 1978.

6. Duket, S. D.; Wortman, D. B.; Seifert, D. J.; Hann, R. L.; and Chubb, G. P.: Documentation for the SAINT Simulation Program. Aerospace Medical Research Laboratory, Wright-Patterson Air Force Base, Ohio, AMRL-TR-77-63, 1978.

7. Duket, S. D.; Wortman, D. B.; Seifert, D. J.; Hann, R. L.; and Chubb, G. P.: An Example to Illustrate the Use of SPSS for the Analysis of a SAINT Model. Aerospace Medical Research Laboratory, Wright-Patterson Air Force Base, Ohio, AMRL-TR-77-64, 1978.

8. Wortman, D. B.; Duket, S. D.; and Seifert, D. J.: SAINT Simulation of a Remotely Piloted Vehicle/Drone Control Facility: Model Development and Analysis. Aerospace Medical Research Laboratory, Wright-Patterson Air Force Base, Ohio, AMRL-TR-75-118, 1976.

9. Kuperman, G. G.; and Seifert, D. J.: Development of a Computer Simulation Model for Evaluating DAIS Concepts. Proceedings: Human Factors Society 19th Annual Meeting, Dallas, Texas, Oct. 1975, pp. 347-353.

10. Seifert, D. J.: Combined Discrete Network — Continuous Control Modeling of Man-Machine Systems, Aerospace Medical Research Laboratory, Wright-Patterson Air Force Base, Ohio, AMRL-TR-79-34, March 1979.

11. Chubb, Gerald P.; and Berisford, Kathleen M.: Manned System Modeling, SAINT Applied to Strategic Navigation. Record of Proceedings, Tenth Annual Simulation Symposium, 1977, pp. 321-352.

12. Askren, W. B.; Campbell, W. B.; Seifert, D. J.; Hall, T.; Johnson, R. C.; and Sulzen, R. H.: Feasibility of a Computer Simulation Method for Evaluating Human Effects on Nuclear Systems Safety. Air Force Human Resources Laboratory, Brooks Air Force Base, AFHRL-TR-76-18, 1976.

13. Wortman, David B.; Sigal, C. Elliot; Pritsker, A. Alan B.; and Seifert, Deborah J.: New SAINT Concepts and the SAINT II Simulation Program. Aerospace Medical Research Laboratory, Wright-Patterson Air Force Base, Ohio, AMRL-TR-74-119, Section VI, 1975.

14. Maltas, Kenneth L.; and Buck, James R.: Simulation of a Large Man/Machine Process Control System in the Steel Industry. Proceedings: Human Factors Society 19th Annual Meeting, Dallas, Texas, Oct. 1975, pp. 193-205.

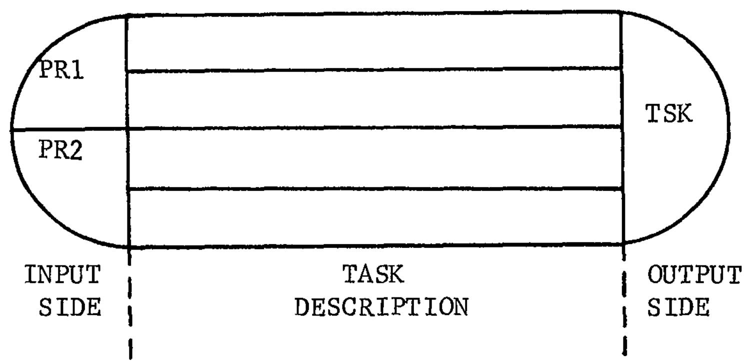
15. Teichner, Warren H.: Human Performance Simulation. Annual Report, AFOSR Contract No. F44620-76-C-0013, Bolling Air Force Base, Washington, D. C., 1976.

16. Wortman, D. B.; Hixson, A. F.; and Jorgensen, C. C.: A SAINT Model of the AN/TSQ-73 Guided Missile Air Defense System. Proceedings of the 1978 Winter Simulation Conference, December, 1978, pp. 879-888.

TABLE I

TASK DESCRIPTION CODES

LABL	task label
TIME	performance time characteristics
MODF	moderator functions
DMOD	distribution modification
RESR	resource requirements
PRTY	priority
INCM	information choice mode
DIFF	different predecessor option
PREC	completion precedence
UTCH	user-defined task characteristics
ATAS	attribute assignments
STAT	statistics
MARK	mark information
TCLR	task clearing
RCLR	resource clearing
SWIT	switching
REGL	regulation



PR1 NUMBER OF PREDECESSOR COMPLETIONS REQUIRED FOR FIRST RELEASE OF THE TASK

PR2 NUMBER OF PREDECESSOR COMPLETIONS REQUIRED FOR SUBSEQUENT RELEASE OF THE TASK

TSK TASK NUMBER

Figure 1.- Task symbol.

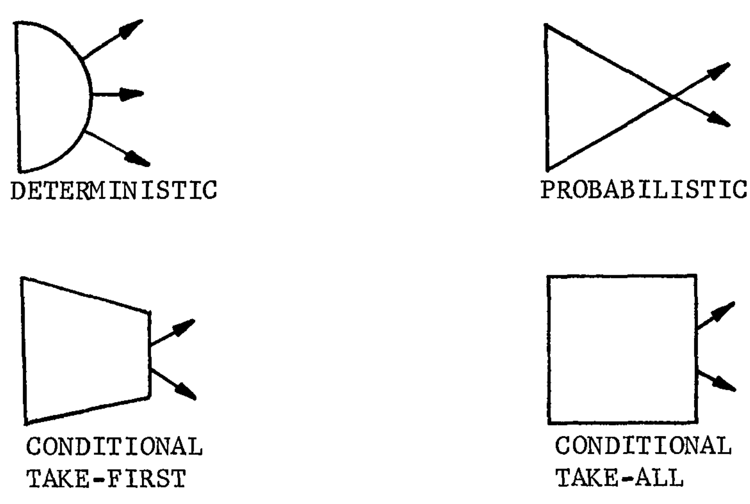


Figure 2.- Task branching symbolism.

ANALYSIS OF VISUAL ESTIMATION OF SYSTEM STATE FROM ARBITRARY DISPLAYS

Patrick A. Gainer
Langley Research Center

SUMMARY

A method is presented for implementing the state estimator of the manual control model when the system output is a visual display of arbitrary form; that is, the display may be pictorial, including real world, or made up of dials and pointers. The method is used to provide error criteria for a look-point controller that appears to be capable of modeling human scanning behavior. This model, if combined with a model of the control process, should be useful in predicting effects of changes in displays on performance of flight tasks.

INTRODUCTION

One of the most important elements of a model of manual control is some form of state estimator. This element receives system outputs and converts them into an estimate of the system state in a form suitable for input to a control algorithm.

The state estimator is usually modelled as a Kalman estimator, which minimizes the variance of the estimated state, and which is capable of accepting sampled data. The output of the estimator includes an estimate of the covariance of the state estimate. This covariance depends on the probable errors in the data, and not on the data. In many cases, such an estimate of covariance is sufficient for an analysis of system performance by the use of covariance propagation techniques. Furthermore, when covariance can be estimated prior to actual data input, an optimum sequence of samples can be predetermined.

Due to the nature of the visual sense, human observers are usually forced to scan a scene in a series of lookpoints in order to extract its information content. If the scene is changing, each lookpoint constitutes a sample of output data of a dynamic system. All that is required for applying the Kalman estimator as a model of human visual observation is a means of estimating probable errors of observation at any lookpoint.

This paper presents a means of estimating the errors to be expected when a human observer estimates the state of a system by looking at a display of some set of system outputs. The display may be pictorial, including "real world," or made up of discrete dials, pointers, etc.

First, the method of estimating the covariance of the state estimate at a given lookpoint will be described. Then the means, and some results, of devising a lookpoint controller to simulate human scanning behavior will be discussed. This latter work is presented in full detail in reference 1.

SYMBOLS

x	state vector
$\hat{x}$	estimated state vector
y	output vector
$[C]$	matrix of coefficients relating state and output vectors
V	standard error of observation of output
$[C]^T$	transpose of C
$[cov(y)]$	covariance matrix of output vector
$[cov(x)]$	covariance matrix of estimated state vector
θ	pitch angle
ϕ	roll angle
ψ	yaw angle

Abbreviation:

VSI	Vertical Speed Indicator
GSI	Glide Slope Indicator

A dot over a variable denotes a derivative with respect to time.

METHOD OF ANALYSIS

Figure 1 is a manual control block diagram that is borrowed from reference 2. The state estimator in figure 1 will be considered to have in it some means of solving the relationship:

$$y(t) = C\,x(t) + V(t) \tag{1}$$

where $x(t)$ is system state vector
C is a matrix of constant sensitivities

$y(t)$ is visible system output vector
$V(t)$ is random error vector

The sources of error $V(t)$ are in the display and measurements $V_m(t)$ and in the visual sensing of the pilot $V_d(t)$.

$$V(t) = V_m(t) + V_d(t)$$

The visual errors V_d are the ones that vary with lookpoint. Thus, figure 1 may be considered to represent the pilot model at a fixed lookpoint. At a different lookpoint, V_d , and therefore y and $\hat{x}$, may be different.

The perceived system output $y(t)$, which is the input to the state estimation process, might be specified in several ways. For example, in a display composed of discrete instruments, one of the elements of $y(t)$ might be taken to be the altitude, since that quantity is displayed by the altimeter. A more general approach that allows treatment of pictorial as well as discrete displays is to take as elements of $y(t)$ the displacement, rate of change of displacement, and rate of change of displacement of points in the display (e.g., points on the altimeter needle). Consider the display to be broken into segments. The portion of the display in each segment is rigid so that one point in a segment may be taken to represent the whole segment. Rotating display elements should be represented by more than one segment, so that rotation of any one segment can be neglected. This point may be seen to move vertically or horizontally or both in response to one or more of the state variables $x(t)$. Taking each component as a separate indicator of system output, each point in the display may provide four elements of $y(t)$: vertical and horizontal displacement and vertical and horizontal rate of change of displacement from a nominal position.

The next step is to perturb each state variable in turn and to calculate the resulting effects on each display segment. These effects will be expressed as linear influence coefficients which are, either exactly or approximately, partial derivatives of vertical and horizontal angular displacement and rate of displacement, as measured at the observer's eye, with respect to the system state variables. These influence coefficients are, of course, the elements of the matrix C in equation 1.

The random error $V(t)$ remains to be specified. That part of it that is due to the visual sense $V_d(t)$ depends on the observer's acuity at each display point, which in turn depends on the location of the display point with respect to the observer's lookpoint. From a knowledge of the observer's resolving power at any point in his visual field and of where he is looking, one can estimate the element of $V(t)$ for each display segment. Typical resolution curves are shown in figures 2 and 3. For any given lookpoint, the eccentricity angle (visual angle between lookpoint and display point) of each display point is calculated, and the corresponding resolutions from figures 2 and 3 are taken as standard deviations for calculating $V_d(t)$.

Now the minimum variance estimate of $x(t)$ may be formulated from equation 1. This estimate is:

$$\hat{x}(t) = \left[C^T [\mathrm{cov}(y)]^{-1} C\right]^{-1} C^T [\mathrm{cov}(y)]^{-1} y(t) \tag{2}$$

The matrix $\mathrm{cov}(y)$ is the covariance of $y(t)$ which, for the case of uncorrelated measurement errors, is a diagonal matrix formed by squaring the elements of $V(t)$. If measurement errors are known to be correlated, there will of course be off-diagonal terms in $\mathrm{cov}(y)$, but equation 2 is still valid. What is required in this paper, as well as $\hat{x}(t)$, is an estimate of its covariance. It may be shown (ref. 3) that the covariance of $x(t)$ is:

$$\mathrm{cov}\ (\hat{x}) = \left[C^T [\mathrm{cov}(y)]^{-1} C\right]^{-1} \tag{3}$$

It is seen that, since for each different lookpoint there is a different matrix $[\mathrm{Cov}(y)]$, the covariance of the estimated state varies with lookpoint.

Example

In order to demonstrate how to calculate $\mathrm{cov}(x)$ for a given lookpoint, a simple example has been concocted. Two hypothetical displays will be compared for each of two lookpoints.

Figure 4 shows two displays, each of which is capable of showing three variables. When all three variables are at zero, both displays look the same. In Display A all three line segments move together as in figure 4(b). In Display B, each segment responds to a different variable, as in figure 4(c). Suppose that the variables presented are pitch, roll and yaw angles, and let the display be viewed from such a distance that movement of 1 degree visual angle represents 1 degree pitch or yaw (according to direction), and 1 degree rotation of the display in its plane represents 1 degree of roll.

The displays are of such size that each line segment subtend 11 degrees of visual angle.

The display area must be divided into discrete areas, a point in each area being taken as the indicator of system output for that area as shown in figure 5. Sensitivities of these points to changes in state variables are calculated in terms of change of visual angle or angular rate per unit change in each state variable. These calculated sensitivities are given in table I. The sensitivity matrix [C] would have in the general case four rows for each segment. For the illustrative cases in this paper, certain movements were considered negligible or not visible. The sensitivities of these movements, being zero, were omitted from the table to save space. For example, horizontal movement, due to rotation, of a point on a horizontal line was neglected. Also, motion along a line was considered not visible. If the lines were really made up of dots, this motion could be seen, and the sensitivity matrix would have corresponding terms. Segment 6 is, in fact, the only one in which both vertical and horizontal components of motion were considered to be visible, and so it appears four times in table I for each display.

The accuracy of observation depends on the observer's acuity and his lookpoint. In order to simplify estimation of parafoveal acuity, it is assumed that contours of constant acuity are circles centered on the lookpoint. This assumption is not essential to the method, but the spread of available acuity data is great enough that a more detailed mapping seems unwarranted.

Two lookpoints were chosen for illustrating the method: one at the intersection of the lines, and one 6 degrees to the right of the intersection, on the horizontal line. For each lookpoint, eccentricities were calculated for the points in the display for which sensitivities were calculated. For these eccentricities, resolutions were read from the curves. The resolutions were used as the elements of [cov(y)].

The covariance matrices of the estimated state cov(x) were computed from equation 3 and presented in tables II and III. Table II is for Display A. Note the way correlation between pitch, roll, and yaw, as shown by the off-diagonal terms, changes with lookpoint. There is no correlation between displacement and rate for either lookpoint.

The covariances for Display B are diagonal matrices for both lookpoints.

As might be expected, the variance of estimate of any given state variable depends on lookpoint. This dependence on lookpoint is much less for those variables that are perceived through sensation of rate, as rate resolution varies much less across the field of vision than does position resolution.

Scanning Behavior

If the display elements did not move while the observer looked around, the covariance of $x(t)$ could be reduced by combining directly the information from several lookpoints. This reduction would be easy to estimate. However, since $\hat{x}(t)$ represents the state of a dynamic system, the observer, in trying to improve his estimate of any state variable by attending to another point in the display, finds that the uncertainty of the information he obtained from the first point increases while he looks at the second. The optimum means of combining sequential observations of a dynamic system is the Kalman estimator. This estimation algorithm also provides a method for deciding which one of a number of possible observations it would be best to make, provided that the probable error of each possible observation is known beforehand. Combining the Kalman estimator with the method of this paper, for estimating the covariance matrix of the output, cov(y), one may devise a lookpoint controller, as has been done in reference 1, from which the following material is taken.

Figure 6 shows the information flow for this controller, which has been applied to the instrument array shown in figure 7. Figure 8 shows what happens to the variances when the lookpoint is arbitrarily forced to follow the time history in the figure. For these results, the system state transition matrix that governs the growth of covariance is taken to be that of a second order dynamic system without damping or cross-coupling.

In order for the lookpoint controller to choose its own lookpoint, it must have some strategy. Figure 9 illustrates what happens when the lookpoint is chosen so as to provide information about the state variable with the greatest weighted variance. (It is necessary to weight the variances because of dimensionality of each state variable and its importance in control of the system.) Figure 9 represents a case of an autopilot-controlled ("coupled") landing approach where the command bars in the flight director are inactive. Figure 10 is the computed time history of lookpoints over 6.5 seconds, when the GSI and artificial horizon are combined into a single lookpoint identified as Flight Director.

A manually controlled landing approach was simulated simply by adding command bars as variables that need attention. The same lookpoint selection strategy was used, with results as shown in figure 11 when command bars are included along with GSI and artificial horizon in the Flight Director.

In spite of the many simplifying assumptions, the time histories in figures 10 and 11 are quite "Humanoid." The flight director gets most of the attention, and it gets more attention in manual control than in monitoring the autopilot (68 percent of total time compared to 57 percent in monitoring). During monitoring, transitions between peripheral instruments are more likely to happen than during manual control, where nearly all transitions are between flight director and peripheral instruments. However, because of the assumptions and especially because a number of instruments were omitted (Horizontal Situation Indicator, for example), there is no direct comparison with available eye movement data.

CONCLUDING REMARKS

A method has been presented for implementing the state estimator of the manual control model when the system output is a visual display of arbitrary form; that is, the display may be pictorial, including real world, or made up of dials and pointers. The method has been used to provide error criteria for a lookpoint controller that appears to be capable of modeling human scanning behavior. This model, if combined with a model of the control process, should be useful in predicting effects of changes in displays on performance of flight tasks.

REFERENCES

1. Mout, Michael L., Burgin, George H., and Walsh, Michael J.: Use of Sensitivity Analysis to Predict Pilot Performance as a Function of Different Displays. NASA CR 2906, November 1977.

2. Karmarker, J, S., and Sorensen, J. A.: Information and Display Requirements for Independent Landing Monitors. NASA CR 2687, August 1976.

3. Gainer, Patrick A.: A Method for Computing the Effect of an Additional Observation on a Previous Least Squares Estimate. NASA TN D-1599, January 1963.

TABLE I.- SENSITIVITY COEFFICIENTS

(a) Display A

Segment	State Variable: θ	ϕ	ψ	$\dot{\theta}$	$\dot{\phi}$	$\dot{\psi}$
1	1	-.1738	0	0	0	0
2	1	-.1391	0	0	0	0
3	1	-.1045	0	0	0	0
4	1	-.070	0	0	0	0
5	1	-.035	0	0	0	0
6	1	0	0	0	0	0
6	0	0	1	0	0	0
7	0	.035	0	0	0	0
8	0	.070	0	0	0	0
9	0	.1045	0	0	0	0
10	0	.1391	0	0	0	0
11	0	.1738	0	0	0	0
12	0	.035	1	0	0	0
13	0	.070	1	0	0	0
14	0	.1045	1	0	0	0
15	0	.1391	1	0	0	0
16	0	.1738	1	0	0	0
1	0	0	0	1	-.1738	0
2	0	0	0	1	-.1391	0
3	0	0	0	1	-.1045	0
4	0	0	0	1	-.070	0
5	0	0	0	1	-.035	0
6	0	0	0	1	0	0
6	0	0	0	0	0	1
7	0	0	0	0	.035	0
8	0	0	0	0	.070	0
9	0	0	0	0	.1045	0
10	0	0	0	0	.1391	0
11	0	0	0	0	.1738	0
12	0	0	0	0	.035	1
13	0	0	0	0	.070	1
14	0	0	0	0	.1045	1
15	0	0	0	0	.1391	1
16	0	0	0	0	.1738	1

TABLE I.- Concluded

(b) Display B

Segment	State Variable: θ	ϕ	ψ	$\dot{\theta}$	$\dot{\phi}$	$\dot{\psi}$
1	1	0	0	0	0	0
2	1	0	0	0	0	0
3	1	0	0	0	0	0
4	1	0	0	0	0	0
5	1	0	0	0	0	0
6	1	0	0	0	0	0
6	0	0	1	0	0	0
7	0	.035	0	0	0	0
8	0	.070	0	0	0	0
9	0	.1045	0	0	0	0
10	0	.1391	0	0	0	0
11	0	.1738	0	0	0	0
12	0	0	1	0	0	0
13	0	0	1	0	0	0
14	0	0	1	0	0	0
15	0	0	1	0	0	0
16	0	0	1	0	0	0
1	0	0	0	1	0	0
2	0	0	0	1	0	0
3	0	0	0	1	0	0
4	0	0	0	1	0	0
5	0	0	0	1	0	0
6	0	0	0	1	0	0
6	0	0	0	0	0	1
7	0	0	0	0	.035	0
8	0	0	0	0	.070	0
9	0	0	0	0	.1045	0
10	0	0	0	0	.1391	0
11	0	0	0	0	.1738	0
12	0	0	0	0	0	1
13	0	0	0	0	0	1
14	0	0	0	0	0	1
15	0	0	0	0	0	1
16	0	0	0	0	0	1

TABLE II.- DISPLAY A COVARIANCES

[Units are (arc min)^2 and (arc min/sec)^2]

Lookpoint at Intersection					
θ	$\dot{\theta}$	ϕ	$\dot{\phi}$	ψ	$\dot{\psi}$
.404	0	0	0	0	0
0	488	0	0	0	0
0	0	130	0	-1.66	0
0	0	0	39080	0	-3970
0	0	-1.66	0	.496	0
0	0	0	-3970	0	1489
Lookpoint at 6° to Right of Intersection					
1.146	0	-11.41	0	.824	0
0	508	0	-403	0	34.1
-11.41	0	178	0	-12.8	0
0	-403	0	37146	0	-3143
.824	0	-12.8	0	5.48	0
0	34.1	0	-3143	0	1218

TABLE III.- DISPLAY B COVARIANCES

[Units are (arc min)^2 and (arc min/sec)^2]

Lookpoint at Intersection

θ	$\dot{\theta}$	ϕ	$\dot{\phi}$	ψ	$\dot{\psi}$
.475	0	0	0	0	0
0	886	0	0	0	0
0	0	374	0	0	0
0	0	0	85561	0	0
0	0	0	0	.475	0
0	0	0	0	0	886

Lookpoint at 6° to Right of Intersection

θ	$\dot{\theta}$	ϕ	$\dot{\phi}$	ψ	$\dot{\psi}$
5.98	0	0	0	0	0
0	1006	0	0	0	0
0	0	38.8	0	0	0
0	0	0	75960	0	0
0	0	0	0	4.56	0
0	0	0	0	0	952

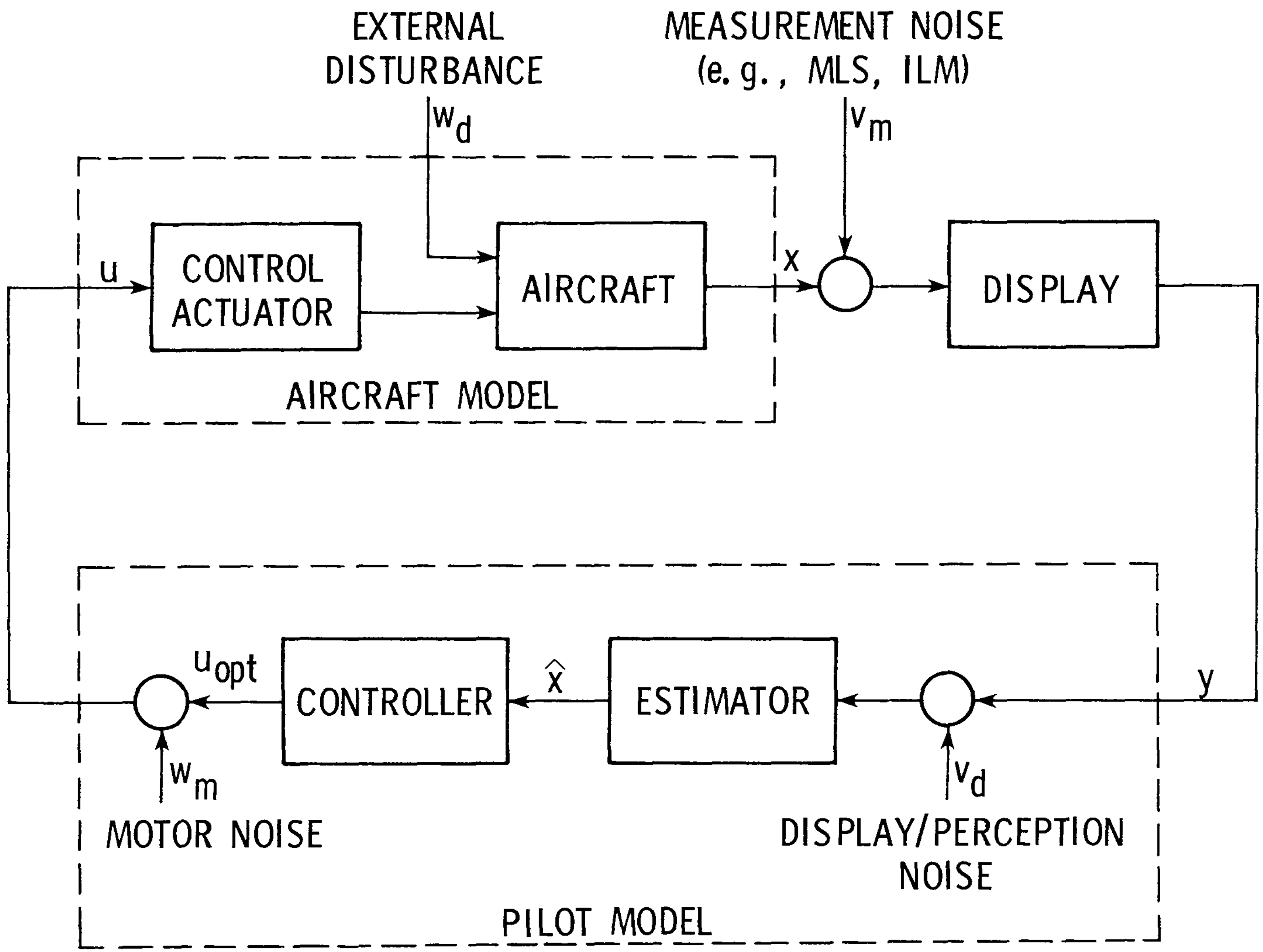


Figure 1.- Manual control system block diagram.

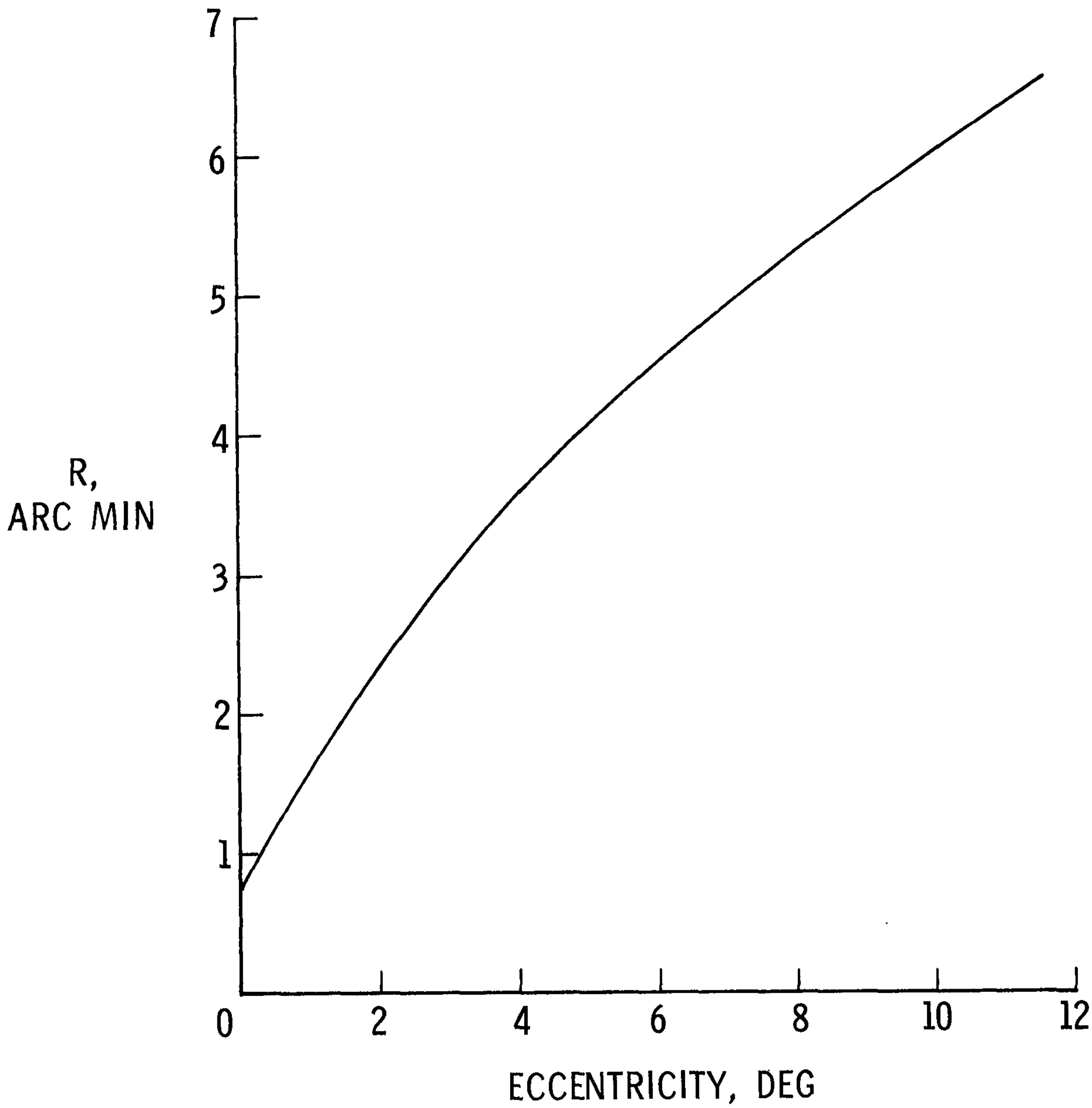


Figure 2.- Displacement resolution versus foveal eccentricity.

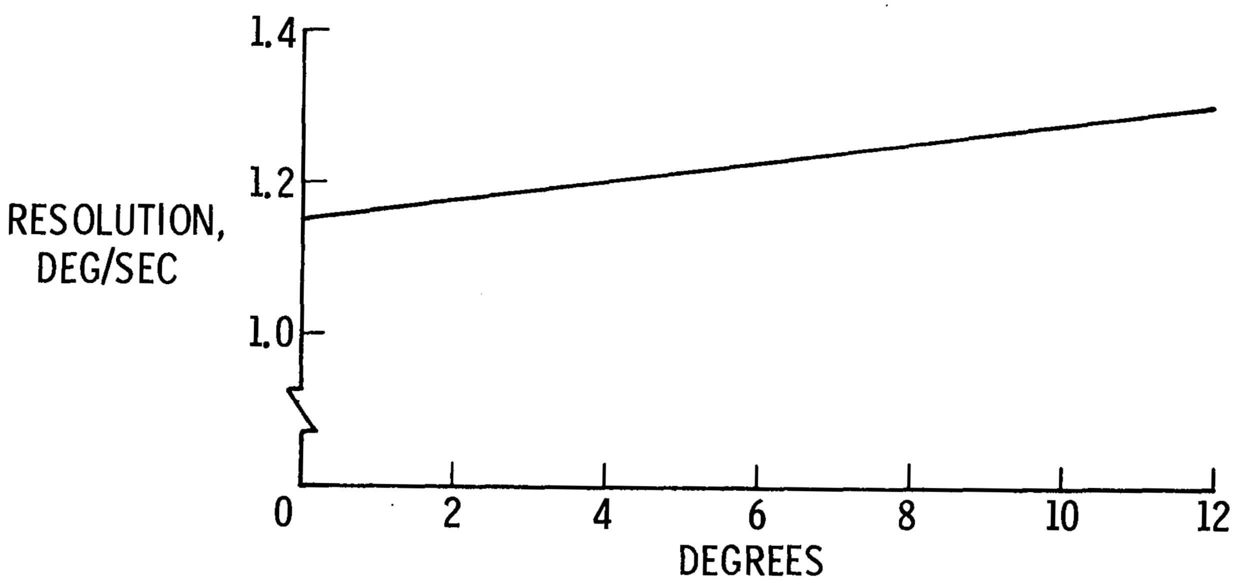
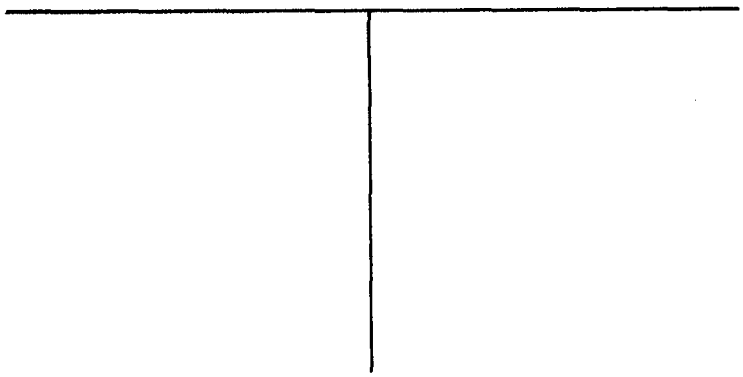
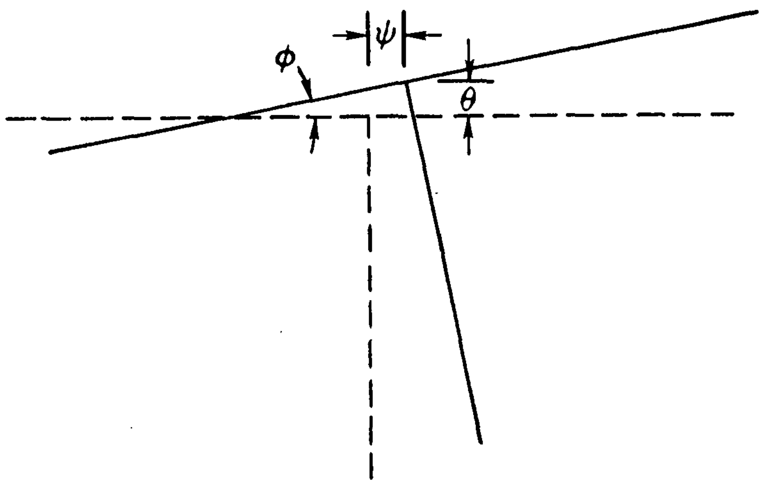


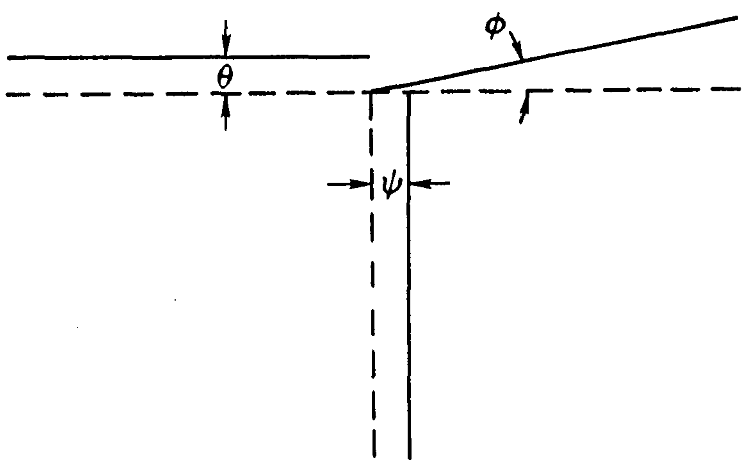
Figure 3.- Rate resolution versus foveal eccentricity.



(a) Display formats A and B at rest.



(b) Display A deflected in pitch, roll, and yaw.



(c) Display B deflected in pitch, roll, and yaw.

Figure 4.- Fictitious displays used in illustrative example.

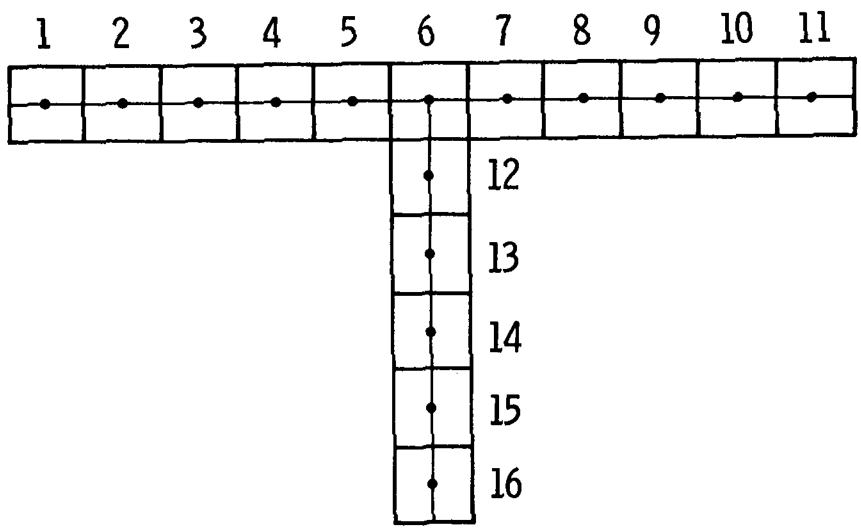


Figure 5.- Display segments.

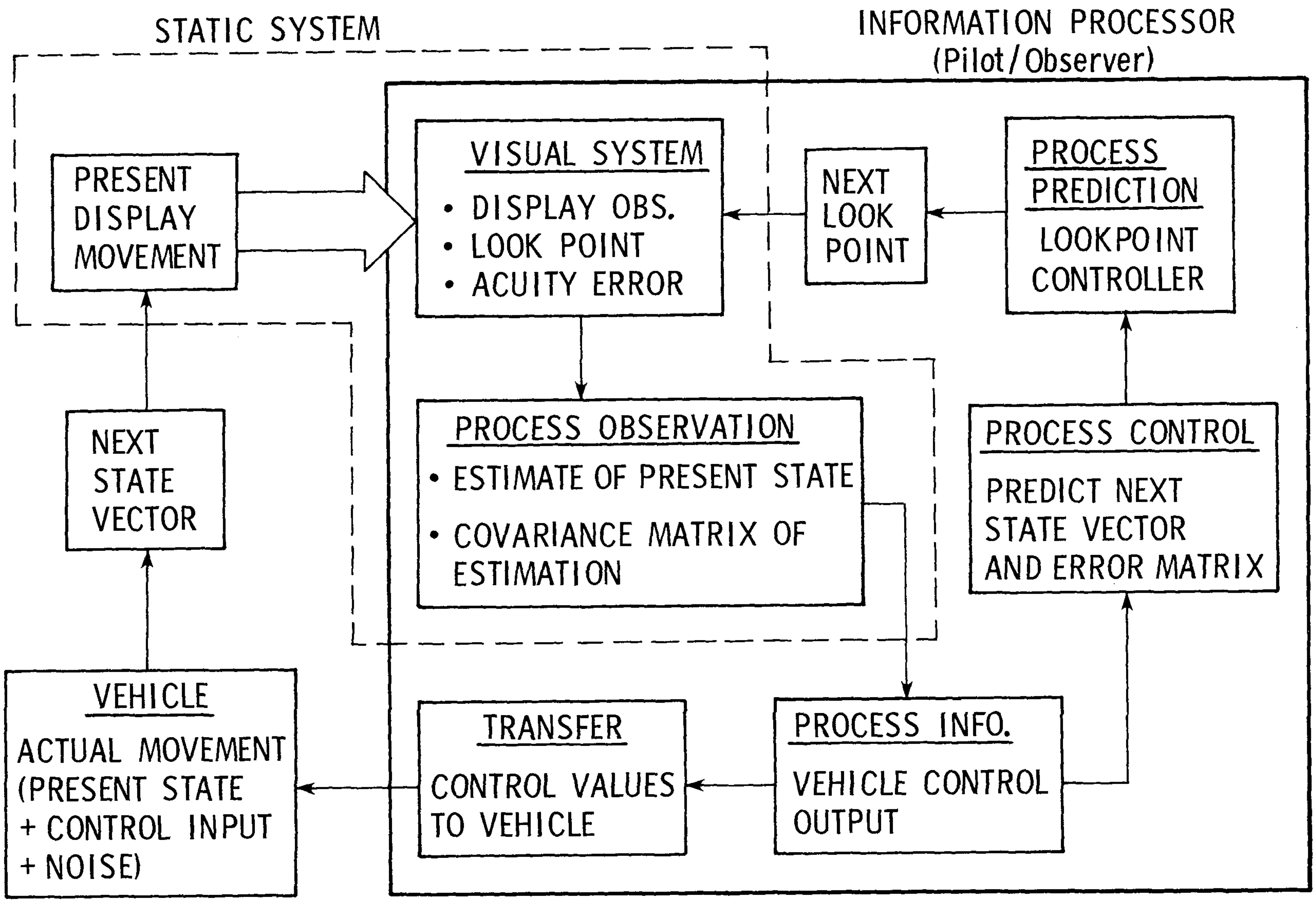


Figure 6.- Information flow for dynamic system.

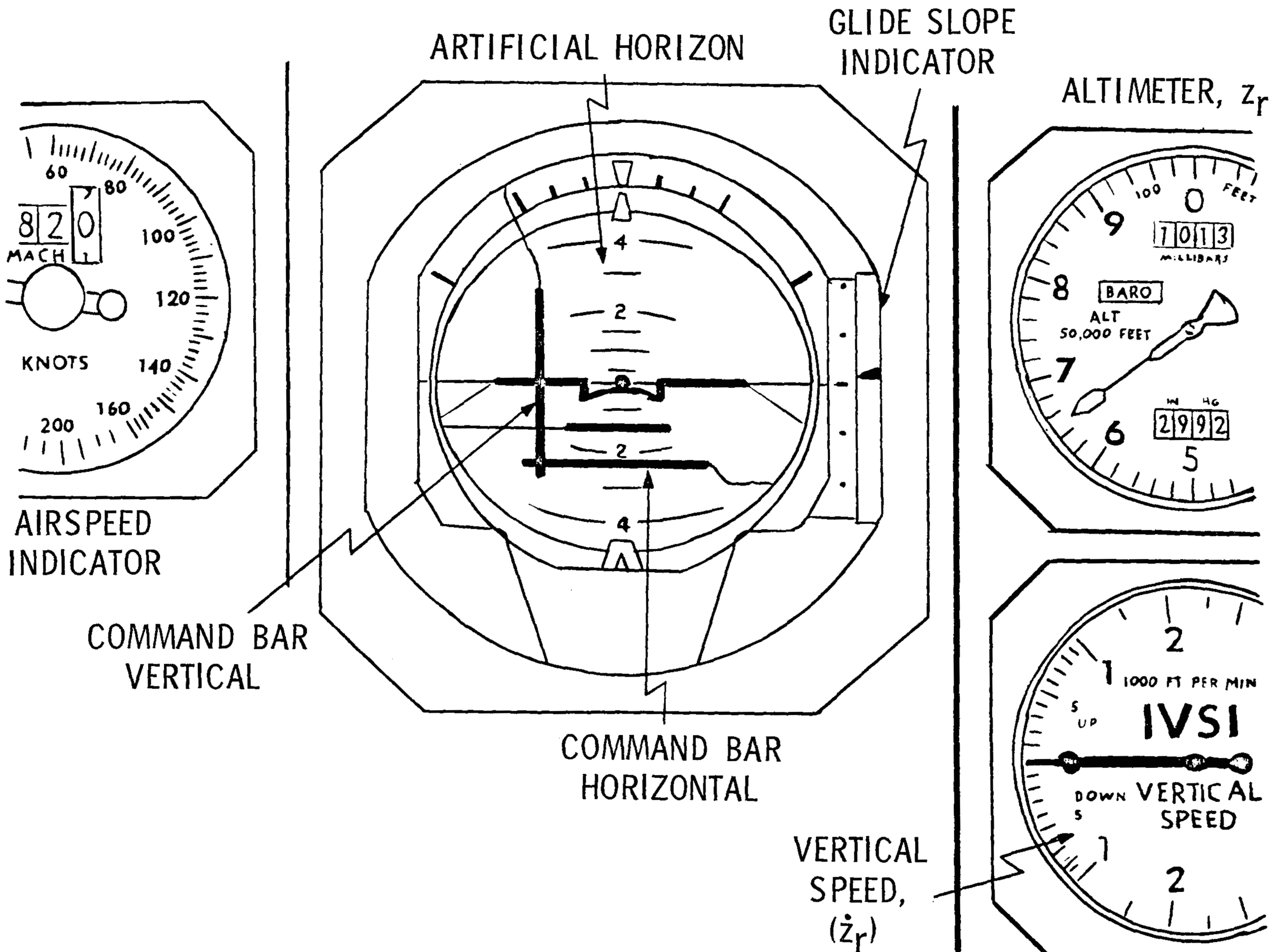


Figure 7.- Display setup.

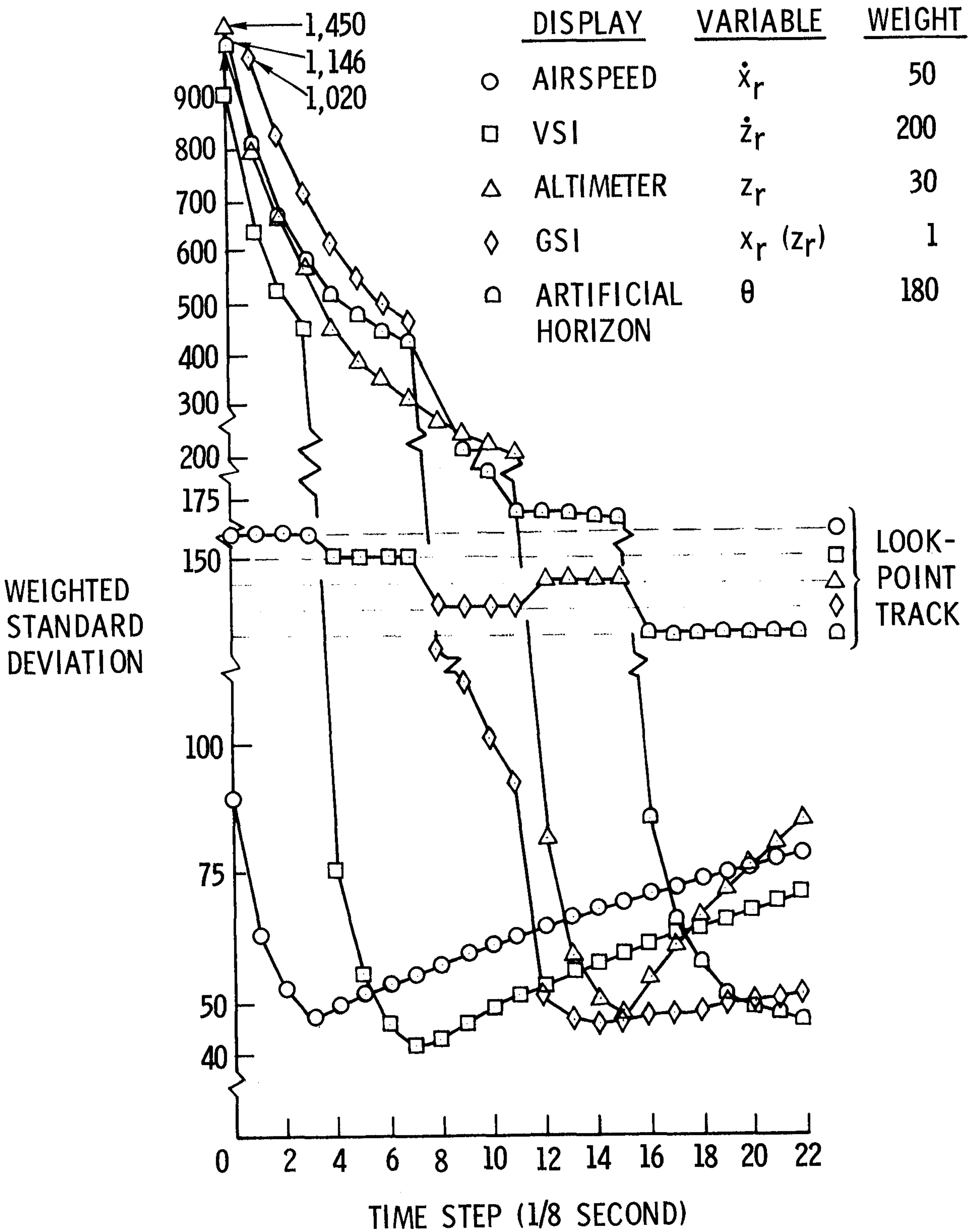


Figure 8.- Fixed lookpoints.

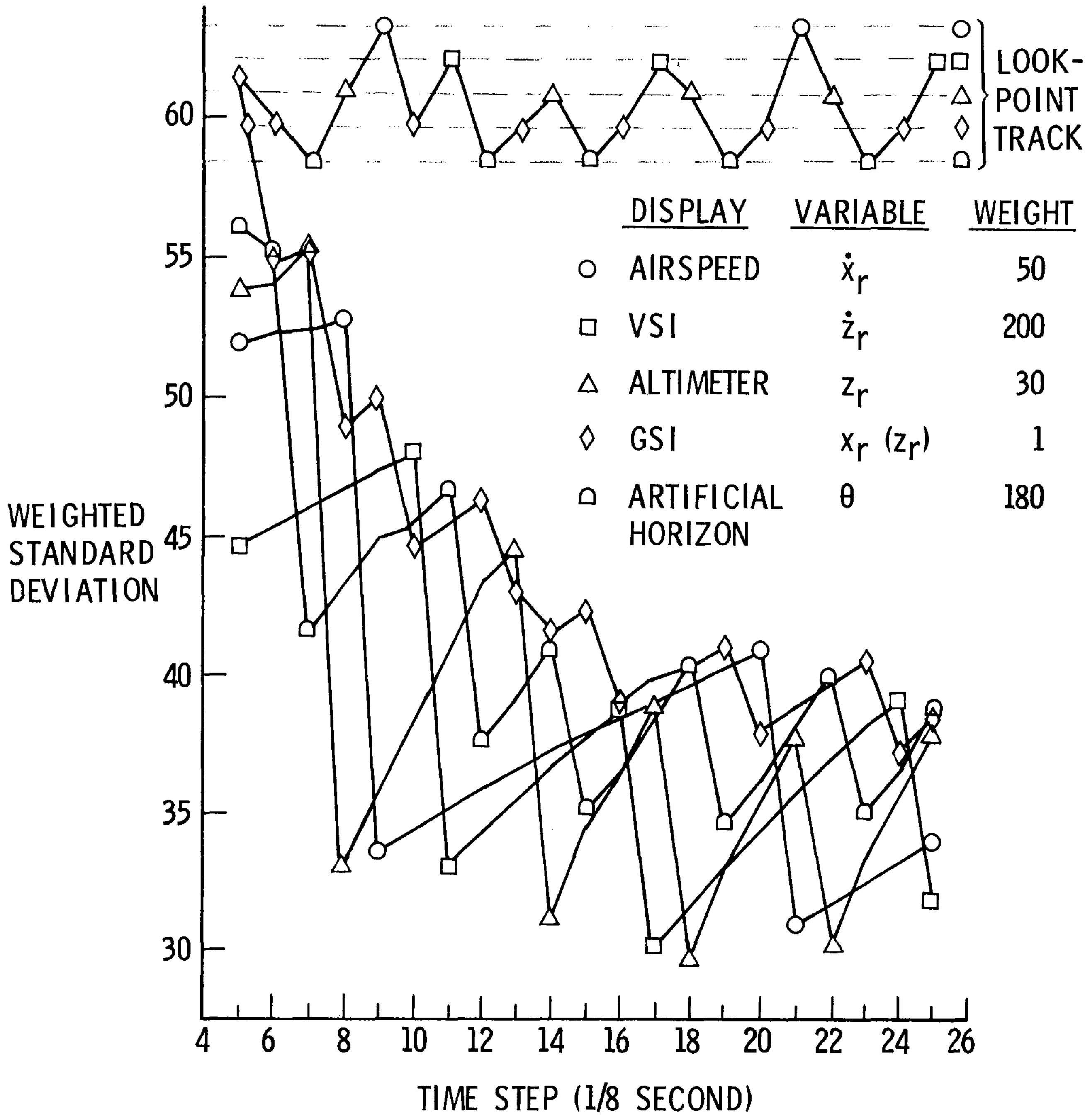


Figure 9.- Lookpoint controller minimizing maximum weighted variance in automatic landing approach.

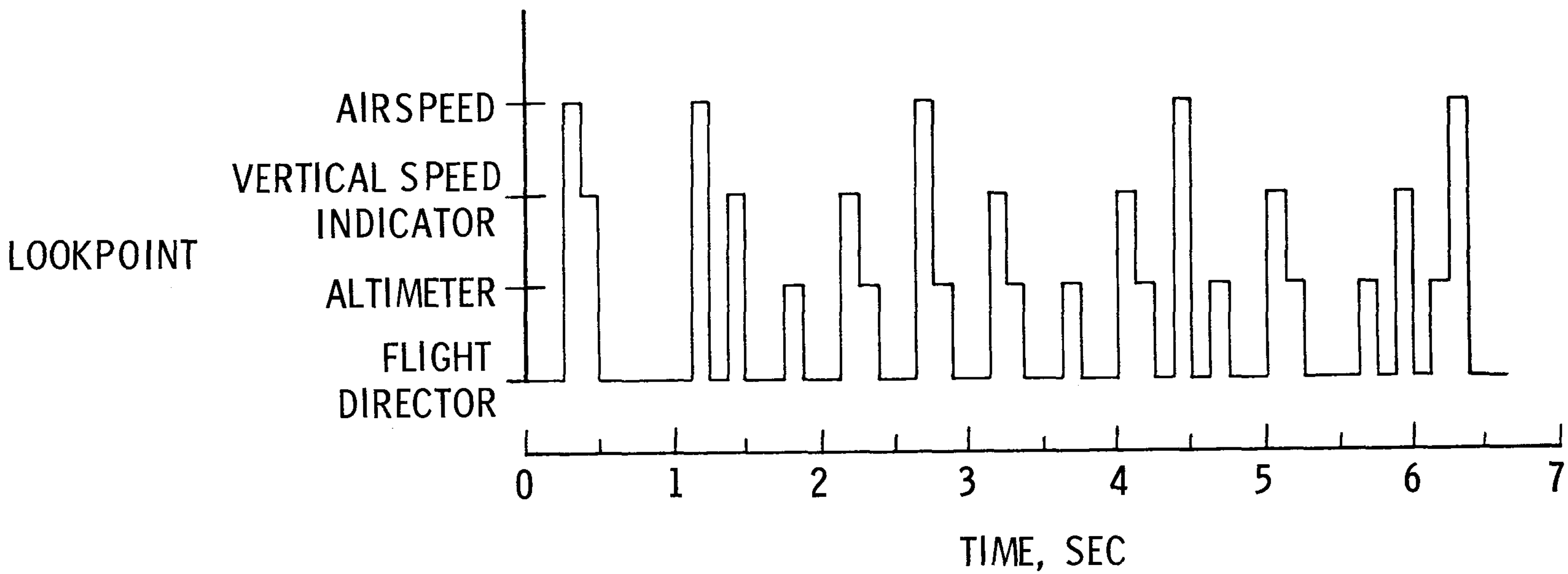


Figure 10.- Time history of lookpoints predicted by model for "coupled" landing approach.

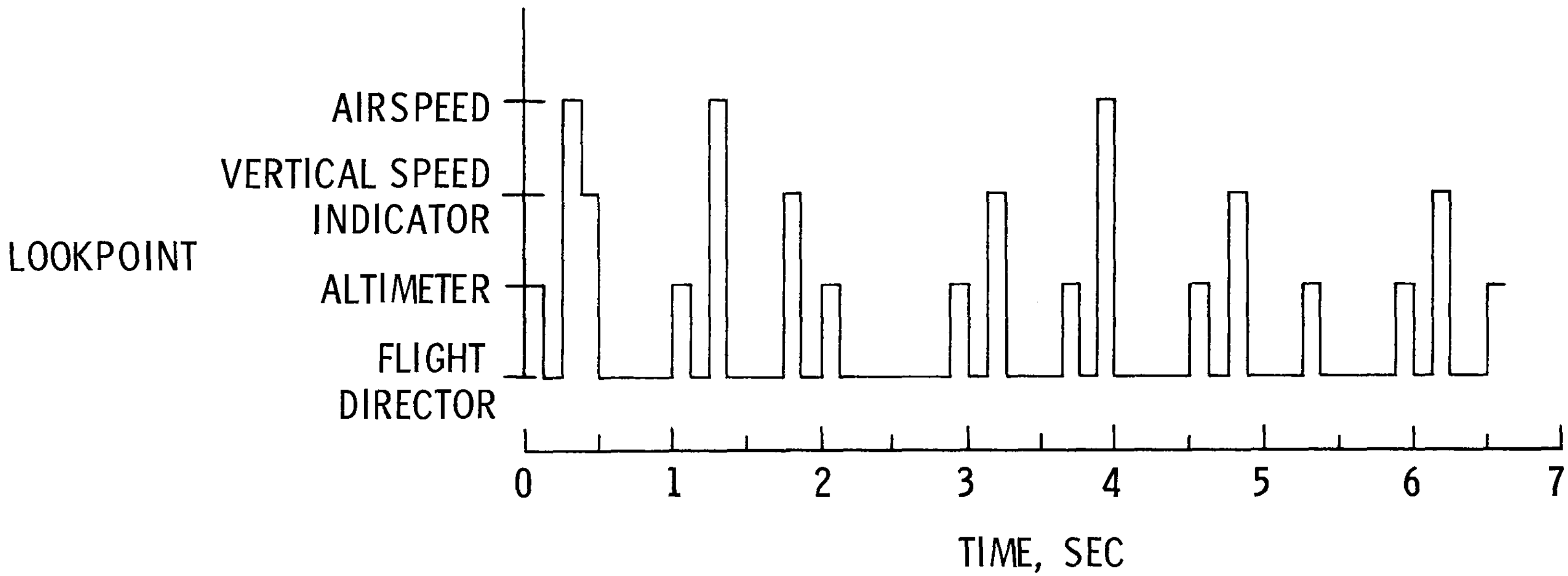


Figure 11.- Time history of lookpoints predicted by model for manual landing approach.

1. Report No. NASA CP-2103	2. Government Accession No.	3. Recipient's Catalog No.
4. Title and Subtitle MODELS OF HUMAN OPERATORS IN VISION DEPENDENT TASK		5. Report Date October 1979
		6. Performing Organization Code
7. Author(s) Marvin C. Waller, Editor		8. Performing Organization Report No. L-13338
9. Performing Organization Name and Address NASA Langley Research Center Hampton, VA 23665		10. Work Unit No. 534-04-13-53
		11. Contract or Grant No.
		13. Type of Report and Period Covered Conference Publication
12. Sponsoring Agency Name and Address National Aeronautics and Space Administration Washington, DC 20546		14. Sponsoring Agency Code

15. Supplementary Notes

16. Abstract

This document is a compilation of several of the presentations included in two tutorial seminars held at the Human Factors Society 1979 Annual Meeting. The discussions present descriptive overviews of some existing models and techniques of modelling the human operator which are at varying stages of completion. Included is a description of HOS, a model developed by Analytics; the Siegel-Wolf model by Applied Psychological Services, Inc.; an Optimal Control Model by Bolt Beranek and Newman, Inc.; and a modelling technique developed for the U.S. Air Force called SAINT. Also a technique of analyzing displays based on a Kalman filter modelling concept of the human operator is included. The latter work is out of NASA Langley Research Center.

17. Key Words (Suggested by Author(s)) Human operator models Pilot models Visual information models	18. Distribution Statement Unclassified - Unlimited Subject Category 53

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price*
Unclassified	Unclassified	86	$6.00

* For sale by the National Technical Information Service, Springfield, Virginia 22161

NASA-Langley, 1979

National Aeronautics and
Space Administration

Washington, D.C.
20546

Official Business

Penalty for Private Use, $300

SPECIAL FOURTH CLASS MAIL
BOOK

Postage and Fees Paid
National Aeronautics and
Space Administration
NASA-451

U.S.MAIL

NASA

POSTMASTER: If Undeliverable (Section 158
Postal Manual) Do Not Return